MODERN
ELECTROCHEMISTRY 2B
ELECTROCHEMISTRY 2B
ELECTROCHEMISTRY 2B

Electrodics in Chemistry, Engineering, Biology, and Environmental Science

SECOND EDITION

JOHN O'M. BOCKRIS AND AMULYA K. N. REDDY

VOLUME 2B

MODERN ELECTROCHEMISTRY

SECOND EDITION

Electrodics in Chemistry, Engineering, Biology, and Environmental Science

This page intentionally left blank

VOLUME 2B

MODERN ELECTROCHEMISTRY

SECOND EDITION

Electrodics in Chemistry, Engineering, Biology, and Environmental Science

John O'M Bockris
Molecular Green Technology
College Station, Texas

and

Amulya K. N. Reddy
President
International Energy Initiative
Bangalore, India

To Tom Bacon

This page intentionally left blank

PREFACE TO THE FIRST EDITION

This book had its nucleus in some lectures given by one of us (J.O'M.B.) in a course on electrochemistry to students of energy conversion at the University of Pennsylvania. It was there that he met a number of people trained in chemistry, physics, biology, metallurgy, and materials science, all of whom wanted to know something about electrochemistry. The concept of writing a book about electrochemistry which could be understood by people with very varied backgrounds was thereby engendered. The lectures were recorded and written up by Dr. Klaus Muller as a 293-page manuscript. At a later stage, A.K.N.R. joined the effort; it was decided to make a fresh start and to write a much more comprehensive text.

Of methods for direct energy conversion, the electrochemical one is the most advanced and seems the most likely to become of considerable practical importance. Thus, conversion to electrochemically powered transportation systems appears to be an important step by means of which the difficulties of air pollution and the effects of an increasing concentration in the atmosphere of carbon dioxide may be met. Corrosion is recognized as having an electrochemical basis. The synthesis of nylon now contains an important electrochemical stage. Some central biological mechanisms have been shown to take place by means of electrochemical reactions. A number of American organizations have recently recommended greatly increased activity in training and research in electrochemistry at universities in the United States. Three new international journals of fundamental electrochemical research were established between 1955 and 1965.

In contrast to this, physical chemists in U.S. universities seem—perhaps partly because of the absence of a modern textbook in English—out of touch with the revolution in fundamental interfacial electrochemistry which has occurred since 1950. The fragments of electrochemistry which are taught in many U.S. universities belong not to the space age of electrochemically powered vehicles, but to the age of

thermodynamics and the horseless carriage; they often consist of Nernst's theory of galvanic cells (1891) together with the theory of Debye and Hückel (1923).

Electrochemistry at present needs several kinds of books. For example, it needs a textbook in which the whole field is discussed at a strong theoretical level. The most pressing need, however, is for a book which outlines the field at a level which can be understood by people entering it from different disciplines who have no previous background in the field but who wish to use modern electrochemical concepts and ideas as a basis for their own work. It is this need which the authors have tried to meet.

The book's aims determine its priorities. In order, these are:

1. Lucidity. The authors have found students who understand advanced courses in quantum mechanics but find difficulty in comprehending a field at whose center lies the quantum mechanics of electron transitions across interfaces. The difficulty is associated, perhaps, with the interdisciplinary character of the material: a background knowledge of physical chemistry is not enough. Material has therefore sometimes been presented in several ways and occasionally the same explanations are repeated in different parts of the book. The language has been made informal and highly explanatory. It retains, sometimes, the lecture style. In this respect, the authors have been influenced by *The Feynman Lectures on Physics.*

2. Honesty. The authors have suffered much themselves from books in which proofs and presentations are not complete. An attempt has been made to include most of the necessary material. Appendices have been often used for the presentation of mathematical derivations which would obtrude too much in the text.

3. Modernity. There developed during the 1950's a great change in emphasis in electrochemistry away from a subject which dealt largely with solutions to one in which the treatment at a molecular level of charge transfer across interfaces dominates. This is the "new electrochemistry," the essentials of which, at an elementary level, the authors have tried to present.

4. Sharp variation is standard. The objective of the authors has been to begin each chapter at a very simple level and to increase the level to one which allows a connecting up to the standard of the specialized monograph. The standard at which subjects are presented has been intentionally variable, depending particularly on the degree to which knowledge of the material appears to be widespread.

5. One theory per phenomenon. The authors intend a *teaching book,* which acts as an introduction to graduate studies. They have tried to present, with due admission of the existing imperfections, a simple version of that model which seemed to them at the time of writing to reproduce the facts most consistently. They have for the most part refrained from presenting the detailed pros and cons of competing models in areas in which the theory is still quite mobile.

In respect to references and further reading: no detailed references to the literature have been presented, in view of the elementary character of the book's contents, and the corresponding fact that it is an introductory book, largely for beginners. In the

"further reading" lists, the policy is to cite papers which are classics in the development of the subject, together with papers of particular interest concerning recent developments, and in particular, reviews of the last few years.

It is hoped that this book will not only be useful to those who wish to work with modern electrochemical ideas in chemistry, physics, biology, materials science, etc., but also to those who wish to begin research on electron transfer at interfaces and associated topics.

The book was written mainly at the Electrochemistry Laboratory in the University of Pennsylvania, and partly at the Indian Institute of Science in Bangalore. Students in the Electrochemistry Laboratory at the University of Pennsylvania were kind enough to give guidance frequently on how they reacted to the clarity of sections written in various experimental styles and approaches. For the last four years, the evolving versions of sections of the book have been used as a partial basis for undergraduate, and some graduate, lectures in electrochemistry in the Chemistry Department of the University.

This page intentionally left blank

PREFACE TO VOLUME 2B

In this volume, the first area to be explored is the photochemistry of processes at the semiconductor/solution interface. The treatment is necessarily quantal and illustrates principles discussed in Chapter 9. Applied to the photosplitting of water, this material has great relevance to future developments in the production, transmission, and storage of energy without adding further to the CO_2 burden of the atmosphere.

The vast edifice of organic chemistry, with its important extensions into pharmacology, has an electrochemical branch that is likely to gain strength as electrocatalysis grows increasingly applicable to organo-electrochemistry. The treatment here has been extremely selective and stresses areas of special interest (e.g., electronically conducting polymers).

The prevention of corrosion is that part of materials science and electrochemistry which, if applied with knowledge, has the potential to save 2–3% of the gross national product, which at present is lost because of the destruction of materials. The field has a strong moving frontier and advantage has been taken of the fact that some of the new information lends itself to diagrammatic presentation.

There is little need to justify the treatment of energy conversion and storage (fuel cell and batteries). The field is in a dynamic phase, having its frontier in methanol-hydrogen fuel cell systems in German, American, and Japanese designs for automobiles; while the use of lithium is greatly advancing the application range of batteries. The newer areas of condenser storers, based upon the large capacitance of the double layer, may yet challenge conventional batteries.

Bioelectrochemistry is hardly a new area—it led to a Nobel prize in the 1950s—but its theory has hitherto been based on older Nernstian principles, and this type of thinking in electrophysiology involves a conservation that slows the introduction of interfacial electrode kinetics in newer treatments. Metabolism, nerve conduction, brain electrochemistry—these areas are where the mechanism of the processes, as yet poorly understood, certainly involve electric currents and are most probably electrochemical.

Last of all is an area that can only grow in importance in the new century: environmental electrochemistry. Here, the leading thoughts concern the greenhouse effect and how the continued use of nonrenewable resources (coal, oil, and natural gas) may have devastating environmental consequences in the next one to two generations. Several of the energy production processes now known to offer healthier alternatives involve electrochemistry, a core subject in this growing field.

Again, each chapter has been reviewed by a specialist in this area. Of course, choices have had to be made as to what areas would be presented. The responsibility for these and also for remaining errors is solely that of the authors.

The problems for this volume follow the scheme outlined in the preface to Vol. 1.

TEXT REFERENCES AND READING LISTS

Because electrochemistry, as in other disciplines, has been built on the foundations established by individual scientists and their collaborators, it is important that the student know who these contributors are. These researchers are mentioned in the text, with the date of their most important work (e.g., Gurney, 1932). This will allow the student to place these leaders in electrochemistry in the development of the field.

Then, at the end of sections is a suggested reading list. The first part of the list consists of some seminal papers, publications which, in the light of history, can be seen to have made important contributions to the buildup of modern electrochemical knowledge. The student will find these earlier papers instructive in comprehending the subject's development. However, there is another reason to encourage the reading of papers written in earlier decades; they are generally easier to understand than the later, necessarily more sophisticated, papers.

Next in the reading list, are recent reviews. Such documents summarize the relevant field and the student will find them invaluable; only it must be remembered that these documents were written for the scientists of their time. Thus, they may prove to be less easy to understand than the text of this book, which is aimed at students in the field.

Finally, the reading lists offer a sampling of some papers of the past decade. These should be understandable by students who have worked through the book and particularly those who have done at least some of the exercises and problems.

There is no one-to-one relation between the names (with dates) that appear in the text and those in the reading list. There will, of course, be some overlap, but the seminal papers are limited to those in the English language, whereas physical electrochemistry has been developed not only in the United Kingdom and the United States, but also strongly in Germany and Russia. Names in the text, on the other hand, are given independently of the working language of the author.

John O'M. Bockris, College Station, Texas
Amalya K. Reddy, Bangalore, India

CONTENTS

CHAPTER 11

SELECTED ASPECTS OF ORGANOELECTROCHEMISTRY

CHAPTER 12

ELECTROCHEMISTRY IN MATERIALS SCIENCE

CHAPTER 13

CONVERSION AND STORAGE OF ELECTROCHEMICAL ENERGY

CHAPTER 14

BIOELECTROCHEMISTRY

CHAPTER 15

ENVIRONMENTALLY ORIENTED ELECTROCHEMISTRY

This page intentionally left blank

CHAPTER 10

PHOTOELECTROCHEMISTRY

10.1 INTRODUCTION

Photoelectrochemistry deals with the electrochemical current produced when one shines a light onto an electrode in solution. Photoelectrochemistry is a good follow-up subject after a chapter on quantum-related electrochemistry because it involves quantal thinking. The most interesting part of Photoelectrochemistry is concerned with the absorption of incoming light in the valence band of a semiconductor (Section 10.3). This results in excitation of electrons in that band to the conduction band, which in a photoanode is followed by an electrochemical interfacial reaction that uses electrons from ions in the interface to fill the holes in the valence band created by the departure of the excited electron.

As already explained (Section 10.3.1), it is the *p*-doped semiconductors that provide cathodic electrons when irradiated with light of sufficient energy; and the *n*-doped type semiconductors that yield the holes to act as photoanodes when semiconductors are used. In cells involving semiconductors, but driven by an outside power source rather than by incident photons, the situation is reversed; the *n*-type emits electrons and the *p*-type receives them.

There are two main attractions of Photoelectrochemistry. The first concerns its fundamentals. The reasoning about the absorption of light involves the energy gap in semiconductors; the incoming light must contain photons energetic enough to pump electrons up from the valence band to the conduction band. Moreover, when semiconductors are used as electrodes in contact with solutions, there is a shift in the energy of the bands near the surface ("the bending of the bands"), which is of relevance in matching the energy levels of electrons or holes with available states of the same energy (empty or filled, respectively) in ions in the solution (radiationless tunneling). All these considerations are similar (but different in detail) to ones dealt with in solid-state physics since the mid-1930s. Thus, the study of Photoelectrochemistry

serves the desirable goal of strengthening the link between solid-state physics and electrochemistry (Gerischer, 1960–1990).

However, there are also attractive applications of photoelectrochemistry. One of these is the photosplitting of water to yield hydrogen and oxygen. As will be shown in Chapter 15, the need to eliminate the injection of CO_2 into the atmosphere may lead to the increasing use of solar energy, but the diurnal variation in this supply means that the energy has to be stored for night use in a clean medium that can be transmitted economically to distant cities, or used at the receiver sites to make methanol from CO_2, the methanol now being the storage medium for hydrogen, which can be released when needed (e.g., by fuel cells) by re-forming the methanol to hydrogen. The transmission of gaseous hydrogen in pipes is cheaper than that of electricity if the distance is more than a few hundred miles. Hydrogen for use in a solar hydrogen economy must be obtained from water if no CO_2 is to be injected into the atmosphere. (If the H_2 is to be released from methanol, the latter should be formed from atmospheric CO_2 to avoid continued accumulation of atmospheric CO_2—for this is evolved in re-forming methanol to the hydrogen needed by fuel cells).

One option is to use photovoltaics to produce electricity from solar light and then electrolyze water with it in a separate plant. However, it may turn out to be cheaper to directly photoelectrolyze water using a *p*-type cathode and an *n*–type anode, each irradiated with solar light, thus producing hydrogen from water directly in one plant.

10.2. MORE ON BAND BENDING AT THE SEMICONDUCTOR/SOLUTION INTERFACE

10.2.1. Introduction

Electrochemical reactions involving semiconductors occur in a more varied way than with metals. For example, if a semiconductor, like *n*-doped silicon, is put in a circuit and used as a cathode in a normal way (e.g., driven by an outside power source), the available electrons come, not from around the Fermi level as with metals, but from the conduction band [Fig. 10.1(a), Fig. 10.2] of the semiconductor. Correspondingly, when one wants to oxidize a redox ion such as Fe^{2+} at *p*-Si by using an outside power source, the electrons emit from the ion in solution in the interfacial region and enter holes in the valence band of *p*-Si. In a metal, they would enter around the Fermi level.

One must recall, here also that the potential–distance relation for a semiconductor (at least, one without a significant concentration of surface states) is qualitatively different from the potential–distance relation at a metal/solution interface. The essential difference is shown in Figs. 10.1 (a) and 10.1(b). In the semiconductor, most of the potential difference at the interface is inside the solid, and only a few millivolts are in the solution. Of course, with the metal, it is mostly in the sudden drop in the double
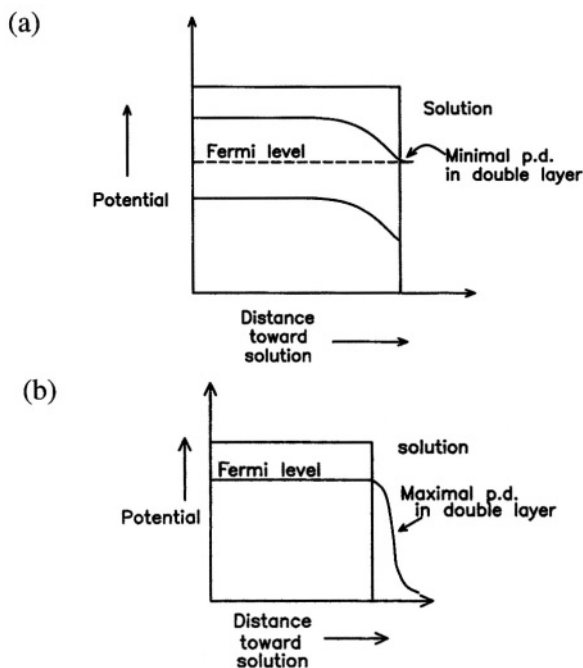
Fig. 10.1. (a) The semiconductor/solution interface (no surface states). (b) The metal/solution interface.

layer, in the metal/solution interface. (There is some drop in potential inside the metal also, but it extends for about 1 Å only.)

10.2.2. Why the Potential Difference in a Semiconductor with No Surface States Is Largely Inside the Solid Phase

At the metal/solution interface, the metal offers little resistance to the flow of electrons. Hence, looking at the matter in terms of Ohm's law (ΔE and IR), the potential difference would be small in correspondence to the small R. Yet when the current has to pass across the interfacial region and be transported no longer by fast-moving electrons, but by the slower moving ions, the resistance is larger—and therefore, so is, the ΔE. So, for a metal/solution interface, we generally neglect the minuscule potential difference in the metal and consider only the potential difference in the interface.

Now, for the semiconductor/solution interface, there are two reasons for the potential difference concentrating *inside* the semiconductor (Fig. 10.1) and being small at the conductivity interface with the solution. On the one hand, the electronic conduction of semiconductors is many orders of magnitude less than that of a
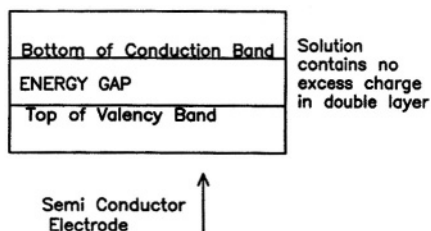
Fig. 10.2. The semiconductor/solution interface in the flatband state. There is no band bending and no excess negative or positive charge on the semiconductor in the interfacial region in a solution-negligible surface state.

metal—hence the R and ΔE, in an ohmic perspective, are orders of magnitude greater. But another reason is that when it comes to the actual interface with the solution, the solid-phase electrons are not concentrated at the surface as with a metal, but are distributed inside the semiconductors (Section 6.6.4). Hence, the excess-charge counter-ions in solution (excluding now the presence of specific adsorption), spread out to the electrons in the solid; there is hardly any Helmholtz layer charge.[1]

One can drop all reference to Ohm's law. Thus, one has (if a metal bears a net negative charge), an excess of positive cations in the diffuse layer in the solution, a charge cloud indeed, and one can find the potential in this diffuse layer (or space charge region) as a function of distance (as it spreads itself over, e.g., 1000 Å). Corresponding to this solution charge cloud (Section 3.3.8), inside the semiconductor one may have an excess negative charge, also extending back into the bulk of the material by perhaps 1000 Å. In fact, the distribution of charge inside the semiconductor (Fig. 10.2) is similar to the distribution of charge due to excess ions in the diffuse double layer in the solution opposite a metal. The equations for the distribution of electrons in the space charge region in the semiconductor are similar to those for the distribution of ions in the solution [see, e.g., Eq. (3.35), Section 3.3.8], but are more realistic because the approximation in the solution treatment that the ions have zero size cannot be too good an approximation for ions, but *is* an excellent approximation for electrons and holes.

10.2.3. Bending the Bands

In semiconductors, the electronic levels occur in bands, energy regions of allowed states that have between them an "energy gap," a forbidden region in which—in the

---

[1]This picture is subject to two caveats. First, in the metal, the potential difference is indeed negligible in the bulk, but directly inside, for about 1 Å, there *is* a potential difference. Correspondingly, some semiconductors will have surface states (Section 10.5.2) and then the Helmholtz layer returns.
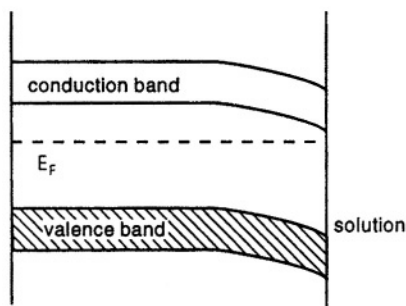
Fig. 10.3. Bending of the bands for an *n*-type semiconductor when the surface potential has become more positive than that of the bulk. (Reprinted from *Interfacial Electrochemistry*, by Wolfgang Schmickler, p. 85, Fig. 7.4a. Copyright ©1996 by Oxford University Press. Used by permission of Oxford University Press, Inc.)

ideal case of semiconductors with no surface states—no electrons or holes exist. One band is characterized by the existence of relatively free electrons (the conduction band). The band at a more negative potential energy, the valence band, is basically full of electrons, but contains some sites in which electrons may be missing ("holes"). In the absence of an excess electric charge on the semiconductor (negative with excess electrons or positive with excess holes), the energy of electrons and holes in these bands becomes independent of distance and one observes, therefore, the *flatband* state (see Fig. 10.2).

Consider now what happens to these flat bands in an *n*-type semiconductor when electrical energy is added to the electrons in them. In contrast to the flatband state with no charge, the energy of the bands changes with distance (i.e., the bands bend) and the change corresponds to the change of the potential in the semiconductor with distance toward the surface. Thus, for an *n*-type semiconductor, if ϕ_S, the potential of the surface, has become (because of the variation of potential with distance) positive with respect to the bulk, the corresponding change in the energy of the bands is to become more negative toward the surface.[2] The positive potential toward the surface attracts electrons there and the surface layer under these conditions becomes an *enrichment layer* (Fig. 10.3).

---

[2] Note the different signs between the energy and electrical potentials. Such a difference arises because of the equation $\Delta G° = -nFE$. Thus, as the electrical potential of the surface becomes electrically more positive, the corresponding potential energy becomes more negative.
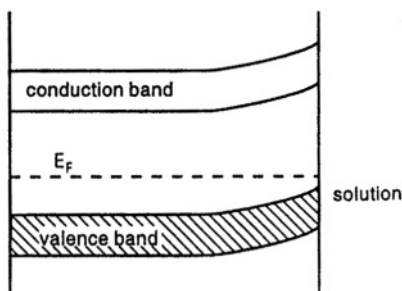
Fig. 10.4. Bending of the bands for a *p*-type semiconductor when the surface is at a more positive potential than the bulk. Here, holes, being positive, are repelled from the positive surface. (Reprinted from W. Schmickler, *Interfacial Electrochemistry* by Wolfgang Schmickler, p. 85, Fig. 7.4d. Copyright ©1996 by Oxford University Press. Used by permission of Oxford University Press, Inc.)

The bands can bend in another direction ("upward") if one makes the potential of the surface more negative than that of the bulk. Because electrons are repelled from it, the surface layer is then called a *depletion layer*. The concentration of electrons in it is less than that in the bulk.

Corresponding but opposite behavior applies to a *p*-type semiconductor (excess of holes on the valence band). A diagram for a *p*-type corresponding to that for an *n*-type with a relatively positive surface is shown in Fig. 10.4.

10.3. PHOTOEXCITATION OF ELECTRONS BY ABSORPTION OF LIGHT

10.3.1. *p*-Type Photocathodes

All photocurrents from a semiconductor, when measured near the limiting current region, have the type of appearance shown in Fig. 10.5. The events that lead to the production, e.g., of hydrogen from the photodecomposition of water on illumination of, say, *p*-type InP, are as follows:

1. The surface of the semiconductor is illuminated and the basic condition in respect to the frequency needed to produce a photocurrent is
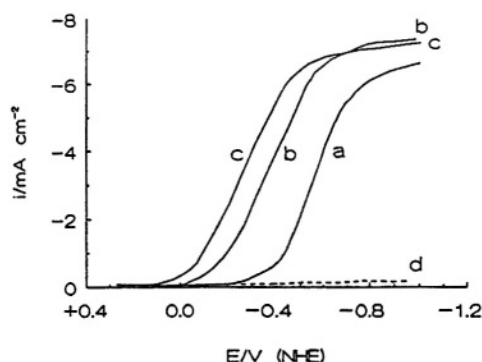
Fig. 10.5. Linear sweep voltammetry of illuminated p-GaAs in 1 M H_2SO_4. (a) Before etching; (b) after etching; (c) after evolving H_2 15 min; scan speed, $v = 100$ mV s^{-1}; (d) under dark conditions. (Reprinted from J.O'M. Bockris, A.Q. Contractor, and S.V.M. Kahn, *Int. J. Hydrogen Energy* **9**: 742, 1984. Reproduced with permission of T. Nejat Veziroglu.)

$$hv > E_g$$

where E_g is the energy gap and v is the frequency of light striking the electrode. This condition means that the energy of the incident photons is such that when a photon arrives in the valence band of the semiconductor, it will have sufficient energy to activate an electron from that band to the conduction band i.e., move it over the energy gap (see Fig. 10.2). Now, a p-doped semiconductor has very few electrons in its conduction band (by contrast, it has many holes in the valence band). The arrival of photoelectrons in the almost empty conduction band will therefore be significant.

2. What happens to an electron newly arrived in the conduction band in a p-type cathode when it is illuminated? Many of the newly arrived electrons will diffuse into the space charge region (see Fig. 10.1) and will be gripped by the electric field that exists there and impelled further to the interface of the semiconductor electrode with the solution.

3. When one refers to the electron as being activated "up" to the conduction band, the vertical movement referred to is on an energy (and not a distance) scale. The subsequent horizontal movement (on a distance scale) of the positively charged hole that the electron has left in the valence band will be away from the surface, i.e., the hole will be impelled by the electric field toward the bulk and thus become physically separated from the electron that formerly occupied it. This reduces the probability of an undesirable event—the deactivation of an electron in the conduction band and its loss to the conduction band as a result of recombination with a hole.

Thus, it must not be thought that all the electrons activated up to the conduction band reach the surface and emit to states in the cations waiting for electrons in the first layer of ions in solution. Much can happen on the way to the surface. The photon penetrates into the semiconductor a distance represented by $1/\alpha$, where α is the absorption coefficient of light θ at a given frequency, before it is absorbed and gives its energy to an electron in the valence band, thus activating it to the conduction band. Then, for those newly arrived photoelectrons that are pulled toward the surface, the following "accidents" can give rise to a loss of the electrons initially activated. (a) Within the semiconductor, electrons newly activated to the conduction band may collide with impurity atoms, which absorb the electron's energy so that it falls back to the valence band, reducing the number of electrons that get converted finally to reduced materials (e.g., finally, H atom molecules), at the surface, (b) Electrons that do reach the surface may recombine with a hole there and be annihilated.

To observe a photocurrent, it is necessary to complete the light-activated reaction at the interface.[3] Thus, e.g., there must be electron ejection from a p-type photocathode to suitable receptor levels in an Fe^{3+} ion in the interfacial region (see Fig. 10.6).

10.3.2. The *n*-Type Photoanode

The action of an n-type photoanode can readily be understood by analogy to the events in a p-type. For n-type photoelectrodes, the absorption of light is governed by a law similar to that for the p-doped electrode; the illuminating light must contain photons the energy of which is greater than the energy gap of the semiconductor material of which the anode consists. However, there is also a fundamental difference here because n-doped semiconductors have many electrons in the conduction band. The significant aspect of the absorption of a photon and formation of an electron-hole pair is the new hole in the valence band, which counts there because there are hardly any holes in the valence band of an n-type semiconductor. It is these photoderived holes, then, that are the active entities in the operation of a photoanode. On the solution side of the interphasial region, there has to be a molecule or an ion, a substance with outer electrons in an energy state that is the same as that in the available holes.[4] That

---

[3]This all concerns a single electrode. The simplest photoelectrochemical *cell* (see Section 1.4.3) consists of an electrode (an *n*-type photoanode, say) that is illuminated and an unilluminated counter-electrode of metal at which an electrode reaction will occur to take up electrons around the circuit from the photoanode. A reference electrode is used to fix the potential of the photoelectrode, and a potentiostat to keep it at a chosen potential.

[4]In an early acquaintance with holes, it may be less easy to conceive of them as entities than as electrons that have been around an entire century. Holes are pseudo-particles. In reality, it is electrons that do the migrating. However, an electron that moves from its site at state A to a new site and starts at B has left A without an occupant. As the electron continues to migrate in the one direction, it can be at once appreciated that it appears to an observer as though a hole is moving in the opposite direction. Further, the "hole" behaves as though it has a charge opposite to that of the electron. Something similar to these ideas has been met in the heuristic theory of liquids called *hole theory,* (Section 5.4.1), where ions are thought of as migrating by "jumping" into gaps in the liquid called *holes*.
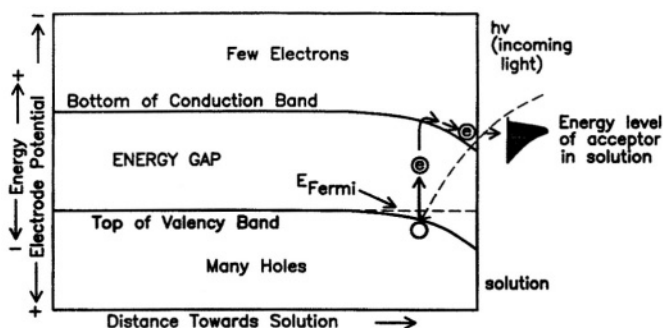
Fig. 10.6. An incoming photon from a source of illumination on a *p*-type electrode makes an electron photoactivated and reaches the conduction band where it electrodiffuses to the surface and emits to energy levels in a suitable ion. The hole created will move toward the bulk and form a photocurrent in the external circuit.

is the condition for radiationless tunneling and that is what will occur to an electron departing from an ion in the interface and reaching a hole in the surface of the semiconductor. All this can be seen in Fig. 10.7.

10.3.3.   The Rate-Determining Step in Photoelectrochemical Reactions

In the mechanisms to be described in this section, one of the idealizations of electrochemistry is being portrayed. Thus, in perfectly polarizable metal electrodes, it is accepted that "no charge passes when the potential is changed." However, in reality, a small current does pass across a "perfectly polarizable" electrode/solution interphase. In the same way, here the statement "free from surface states" (which has been assumed in the account given above) means in reality that the concentration of surface states in certain semiconductors is relatively small, say, less than 10^{12} states cm^{-2}. So when one refers to the low surface state case, as here, one means that the surface of the semiconductor, particularly in respect to sites energetically in the energy gap, is covered with less than the stated number per unit area. A surface absolutely free of electronic states in the surface is an idealization. (If 10^{12} sounds like a large number, it is in fact only about one surface site in a thousand.) A consequence of this is the location of the potential difference at the interphase of a semiconductor with a solution. As shown in Fig. 10.1(a), the potential difference is inside the semiconductor, and outside in the solution there is almost no potential difference at all.

Another characteristic of this kind of semiconductor/solution interface is that a change in potential induced on the electrode from an outside source occurs inside the semiconductor by means of a variation in the Fermi level (see Fig. 10.8). The potential
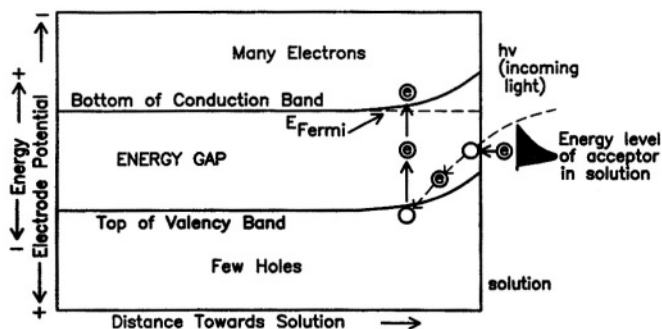
Fig. 10.7. The mechanism of the production of a photocurrent at an *n*-type anode. An incoming photon having an energy > E_g activates an electron in the valence band and this duly goes "up" to the conduction band where it has no marked consequences because of the large number of electrons there. The hole formed at a site (no longer occupied by the electron) is the active entity in the *n*-type semiconductor under the influence of light. It bears a positive charge and hence is attracted to the more negative surface to which it is transported. At the surface (as long as an anion with suitable energy levels is present in the solution in contact with the surface of the *n*-type semiconductor), an electron transfer occurs, and for this process to continue, further holes are transported to the surface and the accepted electron is pulled by the field and the concentration gradient into the bulk of the semiconductor and eventually reaches the outer circuit.

at the surface remains constant when the potential difference across the interphase changes, and the technical term is to say it is *pinned.*

Under these circumstances, then, the rate of the photoelectrochemical reaction is determined by the *rate of transport* of *charge carriers* to the interface. For *p*-type photocathodes, this would refer to the photoactivated electrons that will have been produced from the valence band and are now in the conduction band. The electrons there are impelled to the surface both by diffusion (dependent on the concentration gradient of electrons) and by means of the electric field resulting from the potential gradient near the surface (Fig. 10.8), that depends on the electrode potential that shifts the Fermi level. For a given photoillumination intensity, there will be a fixed limiting current where (at best) the rate of transport to the surface becomes equal to the maximum rate of electron production due to photoactivation (Fig. 10.9). In reality, the numerical value of the limiting current will be also determined by various accidents (e.g., collisional deactivations) that destroy photoelectrons on their way to the interface.
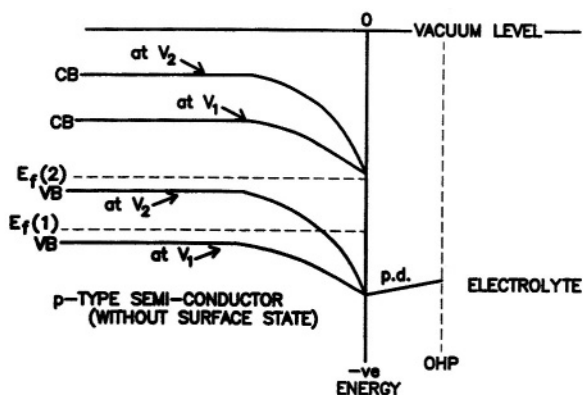
Fig. 10.8. A schematic diagram of a *p*-type semicon-
ductor/solution interface at two applied potentials, V_1
and V_2, in the *absence of surface states*. The diagram
shows the potential drop in the solution's Helmholtz
layer and exhibits no variation in Helmholtz layer po-
tential difference with applied potentials. The Fermi
level is not pinned.

10.3.4. The "Schottky Barrier"

The *n–p* junction was discussed in Section 7.4.1.2. In the original concept, this
junction resulted from the transfer of electrons from one semiconductor to another. In
the figures in Sec. 7.4, potential–distance relations for the junction of *n* and *p*
semiconductors are shown. It is clear that here the transfer of one charge carrier from
one semiconductor to the next in an uphill direction can be thought of as being opposed
by the electrical potential hill shown. Such potential hills are termed "Schottky
barriers."

The same term is sometimes used to describe the potential–distance relations in
semiconductors with a low concentration of surface states (hence the term "Schottky
barrier model"). However, as can be understood by a reconsideration of the mechanism
there (see Figs. 10.6 and 10.7), the so-called "barrier" is either used for the acceleration
of electrons in *p*-type cathodes or the electrodiffusion of holes to the surface in *n*-type
anodes. Nevertheless, the term "barrier" is still applied.

10.3.5. A Theory of the Photocurrent for Semiconductors of Low
Surface State Concentration Near the Limiting Current

Two assumptions were made in the first theory of the photocurrent at irradiated
semiconductor electrodes (Butler,[5] 1977). The first of these was that mentioned in the

---

[5]M. A. Butler, who authored the theory described here, published a whole generation after J. A. V. Butler,
  whose name is attached to the Butler–Volmer equation.

heading. It implies that the value of the photocurrent in the theory is rate determined by the transport of photo-produced carriers inside the semiconductor; that there are no traps in states for electrons at the surface. The second is that as far as the value of the photocurrent is concerned, the important happenings are all *inside* the semiconductor; if there were any surface effects, they were too small to be of significance.

The basis of Butler's reasoning can be seen in the following equation, which refers to the rate of formation of charge carriers (electron, holes) at a distance x from the electrode/solution interface:

$$\frac{\partial p_x}{\partial t} = g_{(x)} + D \frac{\partial^2 p_x}{\partial x^2} - p_x / \tau \tag{10.1}$$

Here, g_x is the rate of generation of charge carriers at a distance x from the electrode surface, which would be proportional to the incident radiation intensity, the absorption coefficient of the light in the semiconductor concerned. The second term on the right comes from Fick's second law (Section 4.2.7) and represents the rate of change in the concentration of charge carriers p_x at x inside the semiconductor caused by diffusion from or to other regions; finally, the last term p_x / τ (where τ is the lifetime of the carriers) allows for the fact that some photoproduced charge carriers decay on the way to the surface (e.g., by collision with impurities on recombination).

The mathematical evolution of the theory is involved, and struggling with its algebra does little to increase our understanding of photoelectrochemistry. The result for a photocurrent provoked by a light with a monochromatic frequency is

$$\frac{i_{ph}\nu}{eI_0} = [L_d + W_0(V - V_{fb})^{1/2}A(h\nu - E_g)^{1/2}] \tag{10.2}$$

where E_g is the energygap; i_{ph} is the current density caused by the absorption of light. ν is the frequency of light of a single wavelength striking the electrode, taken here as monochromatic; V is the term used for "potential," and fb indicates the reference potential of the flat band (Fig. 10.2). There are two other terms: W_0 is the width of the space charge region, the region in which the potential inside the semiconductor is curved (Fig. 10.1); $L_D = \sqrt{D\tau}$, the average distance a charge carrier could diffuse without being felled by some untoward event on its journey from x to the interface; and A is a constant.

Examination of Eq. (10.2) shows that the equation predicts $(i_{ph}h\nu)^2$ linear with $h\nu$, with an intercept of E_g and the slope of this plot ν should yield V_{fb}. Both these quantities should therefore be determinable by application of the equation to measurements of the photocurrent as a function of the electrode potential.

The equation is deduced for diffusion control inside a semiconductor and applies to data on a single semiconductor near the limiting current. However, its application

to a wide variety of semiconductors is not satisfactory unless care is taken to reduce the surface state concentration by avoiding H and O deposition and adsorption (e.g., by the use of anhydrous media; Lewis, 1984), and thus reducing the rate of recombination at the surface. Both McCann (1978) and Kita and Uosaki (1981) observed gross discrepancies between the flatband potential determined by the application of Butler's equation (10.2) to experimental data and those obtained by other methods. Thus, Eq. (10.2) leaves out surface recombination, neglects surface states, and assumes that the rate-determining step is always transport within the semiconductor.

10.4. WHAT HAS BEEN LEARNED ABOUT PHOTOELECTROCHEMISTRY SO FAR?

There are three clear divisions in the photoelectrochemical field. In the first, one shines light upon a metal electrode. Here, the theory is well worked out (Barker, 1974; Khan and Uosaki, 1976), but metals absorb light very poorly compared with semiconductors, and this makes the photocurrents obtained by irradiating them extremely small. The second division concerns the absorption of light by molecules in solution and electron transfer from or to these photoactivated species and to or from a conveniently placed electrode (Albery, 1989). Such phenomena are of interest to photochemists, but here the electrode is the handmaiden of the photochemistry and so we regretfully forgo a description of the material.

The attractive part of the photoelectrochemistry field (particularly if one is interested in the photosplitting of water) relates to the phenomena that occur when light shines on semiconductors in solution. This field awoke from a long slumber in 1972 when Fujishima noticed bubbles evolving as he washed TiO_2 powder under intense light. Soon Fujishima and Honda had built a cell of TiO_2 as the irradiated electrode (it is an n-type semiconductor and so would be a photoanode) and a platinum counter-cathode. They found that H_2 evolved on the platinum when light was shined on the TiO_2; this was the first photoelectrochemical decomposition of water. Around this time (1974), two events occurred that magnified the importance of the discovery. The first was an Arab–Israeli conflict that led to a large increase in the wholesale price of oil; the second was the launching of the idea of clean hydrogen as a substitute for natural gas and gasoline and the fuel for fuel cells (Chapter 13). Could the hydrogen be obtained more cheaply by shining solar light on semiconductor electrodes in electrochemical cells (one plant only), rather than making electricity by solar irradiation of photovoltaic cells and then using this electricity in some other plant to electrolyze water?

The renewed interest in photoelectrochemistry was not only related to the grand scheme of solving the problem of planetary warming by substituting electricity and hydrogen from water for gasoline and natural gas. What of the use of redox systems in solution in contact with irradiated semiconductor electrodes to produce electricity from abundant solar light? What of carrying out commercially important organic

reactions [e.g., photo-oxidizing toxic wastes (Krautler and Bard, 1978)][6], or carrying out photoelectrochemically the well-known Kolbe reaction (Chapter 11) (Bard 1980).

The theory of that part of photoelectrochemistry that involves semiconductors divides itself into two parts. The first part, which is given above, is based on two simplifying assumptions. One is that the region of the i_{ph}–V plot of interest is sufficiently near the limiting current to assume that rate control is determined by the transport of charge carriers inside the semiconductor.[7] The second assumption is certainly simplifying, but how often is it applicable? It is that all the activities involved in producing a photocurrent occur inside the semiconductor; the surface has nothing to do with it. These two assumptions were both part of the Butler theory outlined in the last section.

Now, as to the actual processes in obtaining a photoelectrochemical current at semiconductor electrodes with low surface state concentrations so that the above assumption might apply, the first thing to consider is the energy gap of the semiconductor electrode material. In Fig. 10.10 we show the energy gap of a number of frequently used semiconductors.

It can be seen that the highest intensity of solar light (number of solar photons per unit area and time) occurs with photon energies of about 2.6 eV. Thus, a semiconductor with an E_g of 2.6 eV will absorb all the photons that have an energy greater than this energy. Thus (Fig. 10.10), more than half the solar photons will be absorbed by such an electrode, and the ones not absorbed are of decreasing value as they become smaller in $h\nu$, i.e., in energy.

Suppose one considers semiconductors that have an energy gap higher than 2.6 eV? Then only photons with an energy greater than this (e.g., 3.0 eV) will be absorbed. Consideration of Fig. 10.11 shows that as the energy gap of a semiconductor increases above 2.6 eV, there are fewer and fewer solar photons, although the ones absorbed will have relatively large amounts of energy per photon as one moves toward the UV side of the solar spectrum.

What of semiconductors with E_g's below 2.6 eV (0.8 μm), e.g., say a semiconductor with 1.0 eV (1.3 μm)? Such semiconductors will absorb all the photons having an energy greater than this, which is apparently most of the solar spectrum (see Fig. 10.11). At first, it might seem that the best semiconductors for absorbing solar light are those with the smallest energy gaps (because they would seem to absorb more and

---

[6] In fact, in 1995, a group of well-known photoelectrochemists, including Lewis (United States), Tributsch (Germany), and Uosaki (Japan), published a joint comment on photoelectrochemistry, extolling the lively character of the field.

[7] Of course, all electrochemical reactions, if run fast enough, run into a current maximum and with few exceptions, this is due to a holdup in transport, e.g., for the metal/solution interface, it is the charge-carrying ions in solution that have a maximum transport rate to the electrode. It is somewhat misleading to concentrate one's attention on just the "top bit" of the photocurrent potential graph when transport control comes into play, though certainly it is of interest to find how fast the reaction may proceed. More recently it has been found that at photocurrent densities sufficiently below the transport-controlled limiting current, photoelectrochemical currents obey the usual Tafel equation exhibited by other electrochemical reactions.
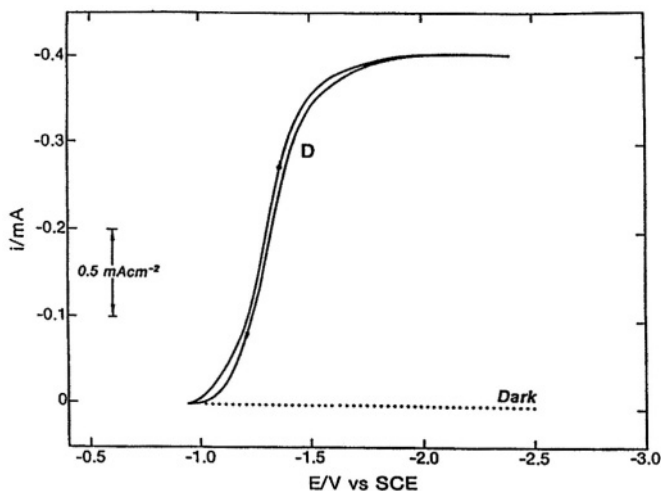
Fig. 10.9. Current–potential curves at a p-CdTe electrode in a DMF solution containing 5% water under irradiation with monochromatic light of 600 nm. CO_2 in a 0.1 M TBAP was reduced largely to HCOOH. (J. O'M. Bockris and J. Wass, *J. Electrochem. Soc.* **136**: 2512, 1989, Fig. 1. Reproduced by permission of the Electrochemical Society.)
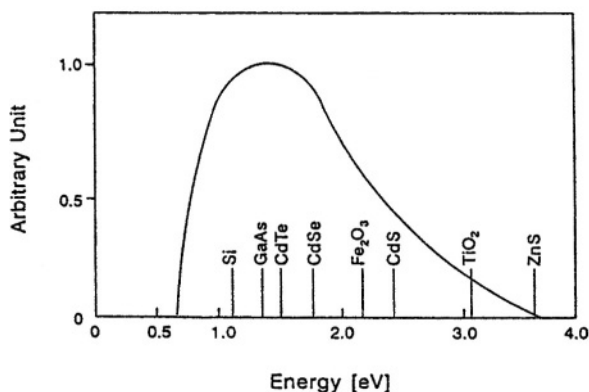
Fig. 10.10. Degree of solar energy absorbed by semiconductors of various band-gap energies. This can be calculated from knowledge of the energy in the solar spectrum at any wavelength and then by taking into account the intensity of light of that energy and the fact that a given semiconductor will absorb only those photons that have an energy greater than its band gap. (Reprinted from A. Gonzalez-Martin, thesis, Texas A&M University 1993, p. 13.)
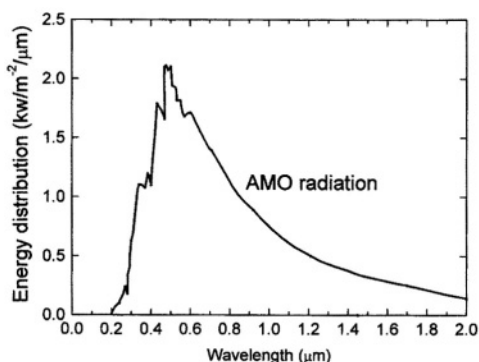
Fig. 10.11. The solar spectrum. AMO indicates that the graph refers to conditions outside the normal atmosphere. (Taken on earth, the spectrum is more complicated because it is indented at wavelengths corresponding to absorption by constituents of the atmosphere.) Note that the relation between the wavelength in micrometers and the energy in electron results is 1.3 eV ≡ μm. Thus, 0.2 μm ≡ 6eV and 2.0 μm ≡ 0.6 eV. As shown in Fig. 10.10, maximum absorption is attained by the semiconductor at ~1.3 eV or 1.0 μm. (Reprinted from M. Green, *Solar Cells*, p. 4, Prentice-Hall, 1984. Reproduced with permission.)

more of the solar spectrum). However, there are counter considerations. For one thing, as can be seen (Fig. 10.11) from the variation in the intensity of the light from the sun with wavelength), the intensity of photons of a given wavelength decreases as the wavelength increases (i.e., as the energy per photon decreases). Thus, the less energetic photons are not worth much individually, and there are not many of them. However, there is a more pressing reason for neglecting materials with low energy gaps. The efficiency of photoelectrochemical conversion depends on how many of the photoactivated electrons make it from the point inside the semiconductor at which they are activated (for p-type photoelectrons) to the interface with the solution. The smaller the energy gap, the easier it is for electrons to deactivate and fall back into the valence band. An optimal energy gap is likely to be an intermediate one. It will absorb the valuable photons (of which there are many), and deactivation will be less easy than that for semiconductor electrode materials with smaller energy gaps.

After successful absorption of photons with energies greater than the energy gap, the events on the way to producing a photocurrent differ for p-type photocathodes and
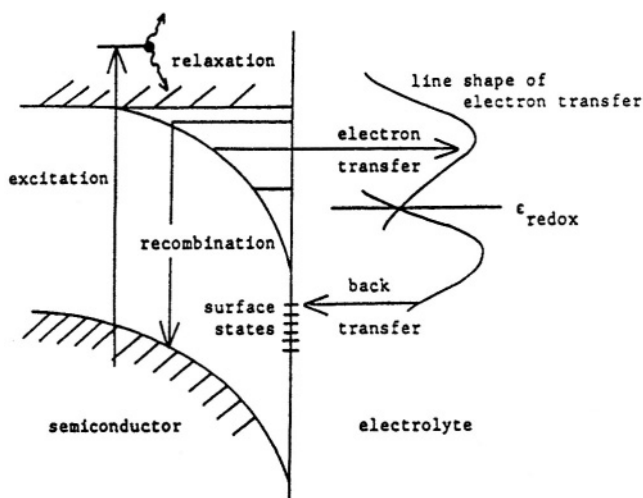
Fig. 10.12. Electronic processes occurring at a semiconductor/electrolyte junction under illumination (Reprinted from E. Buhks, "A New Theoretical Approach to Photoelectron Transfer Across Semiconductor-Electrolyte Interfaces," in *Photoelectrochemistry: Fundamental Processes and Measurement Techniques*, W. L. Wallace, A. Nozik, and S. K. Debb, eds., *Proc. Electrochem. Soc.* 82-3: 13, 1982, Fig. 1. Reproduced by permission of the Electrochemical Society, Inc.)

n-type photoanodes. The details are given above (Figs. 10.6 and 10.7). The p-type cathodes eventually emit photoelectrons from the conductivity band to waiting energy states in ions in solution. For n-type photoanodes, electrons in states in ions in solution that are in contact with the electrode transfer to waiting holes in the electrode and thus enter the valence band (see Fig. 10.12). For the mechanisms described so far—diffusion of electrons or holes inside the semiconductor as the rate-determining step—there are no surface effects (Gerischer, 1975), and hence no electrocatalysis.[8]

---

[8] Studies of systems that minimize surface states and maximize the applicability of theories that are meant to work without them, have become the special province of Prof. Nathan Lewis at the California Institute of Technology, who leads the most active photoelectrochemical research team in the United States. Prof. Lewis started his career with an M.S. in inorganic chemistry at the same famous institute in which he is now a professor, and completed his Ph.D. at MIT in electrochemistry. He represents the archetypal U.S. science professor—aggressive, supreacticve, with many honors and awards, including the ACS award in pure chemistry in 1991. He follows a pattern set by many professorial colleagues of high standing in leading a company that is set to exploit discoveries made in some of his semiconductor research. However, extracurricular activities do not detract from the main characteristic of the research products of the team Lewis leads: fundamental studies of great elegance on systems chosen to diminish surface effects.

10.5. SURFACE EFFECTS IN PHOTOELECTROCHEMISTRY

10.5.1. Introduction

The great emphasis laid by the early workers in the theory of photoelectro-chemistry on processes within the semiconductor arose because the principal authors (Gerischer, 1970s; Pleskov and Gurevich, 1980s) stressed only the region of photocurrent near the limiting current. Hence, the electrode was in a condition in which the transport of electrons or holes within the semiconductor was indeed the rate-determining step. It was Uosaki[9] and Kita (1981) who first found *normal Tafel relations* in photoelectrochemical reactions occurring at current densities well below the limiting-current region. Such a result suggested that interfacial electron transfer is the rate-determining step in photoelectrochemical reactions (Fig. 10.13).

More striking evidence of the effect of surface properties on the rate of photoelectrochemical reactions is provided from experiments on the effect of etching the surfaces of *p*-silicon and evolving H_2 photoelectrochemically on varying surfaces (Szklarczyk and Bockris, 1984). Thus, depending on the degree of etching (i.e., the degree to which an oxide film is removed and the underlying Si exposed the overpotential to reach a certain current density (e.g., 10 mA cm^{-1}) was decreased. The effects are large (see Fig. 10.14); the photocurrent density at a given overpotential changed by about 10^6 times as the structure of the anodic surfaces on *p*-Si was changed.

Thus, below the limiting-current density, for photoelectrochemical hydrogen or oxygen evolution reactions, an interfacial charge-transfer process is rate controlling. In support of this contention, the photocurrent density at a given electrode depends greatly on the nature of the solute species that supplies or takes away electrons at the interface (Gonzalez-Martin, 1993). This would of course be very difficult to interpret if the Schottky barrier approximation (diffusion of internal

---

[9]Uosaki started his career on the scientific staff of Mitsubishi. He came to the Flinders University of South Australia to do his Ph.D. in 1974 and was known among the graduate students there as "the locomotive" for his determination and drive.

He had a productive stay in the Department of Inorganic Chemistry at the University of Oxford carrying out bioelectrochemical work with H. A. O. Hill. In fact, some think that he introduced electrode kinetics to Hill's department (see the latter's contribution to the redox behavior of proteins at metal/solution interfaces, Chapter 14).

Returning to Japan, Uosaki was soon appointed to the prestigious chair of chemistry at the University of Hokkaido, a position formerly held by Juro Horiuti, the most famous of all physical chemists in Japan for his work on catalysis. At Hokkaido, Uosaki is the acknowledged leading physical electrochemist in Japan and has accomplished many things in photoelectrochemistry: the first theory of photoelectrochemical kinetics with a high degree of surface states; the first mathematical treatment (with Kita) of the relation between Schottky-dominated vs. Helmholtz-dominated photoelectrochemical kinetics, and the study of several photostimulated organic reactions, some involving the synthesis of Pharmaceuticals.
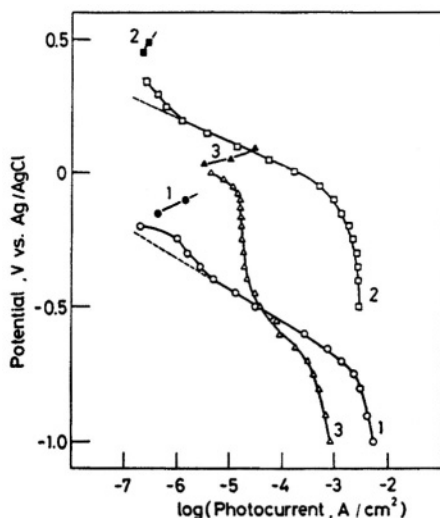
Fig. 10.13. Tafel plots of the photocurrents of *p*-GaP in various solutions. Curve 1, 1 *M* NaOH; curve 2, 0.5 *M* H$_2$SO$_4$; curve 3, 1 *M* NaOH + 1.7 m*M* methyl viologen (Reprinted from K. Uosaki, and H. Kita "Mechanistic Study of Photoelectrochemical Reactions at a p-Gap Electrode," *J. Electrochem. Soc.* **128**: 2156, 1981, Fig. 4. Reproduced by permission of the Electrochemical Society, Inc.)

charge carriers that are in rate control) continued to be rate determining[10] below the limiting-current region.

   Does this mean that the material in the first part of this chapter (where transport of carriers, as in Butler's theory, is assumed to be rate determining) is only applicable near the limiting current? It depends! Near the limiting current for photoelectrochemical reactions, carrier transport is certainly rate determining.[11] As far as the rate control

---

[10]Why should one bother to identify the rate-determining step? One reason is photoelectrocatalysis. If the surface is rate controlling, catalysis may enter the picture and diminish the overpotential needed to obtain a certain current density. The potential available from the light-driven cell becomes greater.

[11]The difference to a limiting current density for the metal–solution case is clear. Thus, for the latter case, the entities, the diffusion of which is limiting in supplying charge carriers near the limiting current, are ions in solution. In the photo case, it is electrons or holes inside the semiconductor that control the supply of charge to the interface. Thus, the limiting currents in photoelectrochemistry are independent of stirring in the solution and proportional to the intensity of illumination, i.e., to the number of minority carriers produced per unit of time.
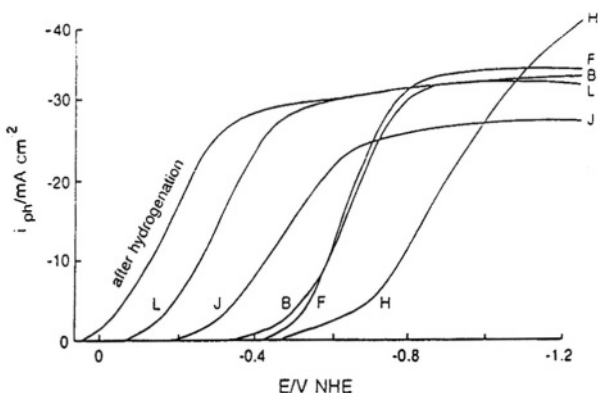
Fig. 10.14. Potentiodynamic runs of *p*-Si electrode in 0.5 *M* H$_2$SO$_4$. Illumination by a 50-mW cm^{-2} Xe lamp. The etching procedures were in HF and HNO$_3$, and from H to L the degree of HF etching increased. (Reprinted from M. Szklarczyk, J. O'M. Bockris, V. Brusic, and G. Sparrow, *Int. J. Hydrogen Energy*, **9**: 707, Copyright 1984. Reproduced with permission of T.Nejat Veziroglu.)

far below the limiting current density is concerned, it may be (if the interfacial reaction is very fast) that the rate control continues to be by means of charge carrier transport. If the surface is relatively free of quantum states (often traps) for electrons and holes, the so-called Schottky case may well be applicable, as shown by the work of N. Lewis (1990s) in nonaqueous solutions where the presence of adsorbed H or O (which would cause such surface states) is unlikely.

However, it seems that the extension of the Schottky (diffusion-controlled) case to the situation below the limiting current is the exception rather than the rule, at least in aqueous solutions. It also depends upon what reaction is being photo-driven. If it is a redox reaction (i.e., the only step in the interfacial reaction is a single electron transfer between the electrode and the redox ions), surface states are liable to be smaller in concentration and the Schottky barrier case—transport within the semiconductor is rate determining—may be more applicable. However, if the photoelectrochemical reaction involves water splitting, and H or O is adsorbed on the electrode surfaces, intermediates of adsorbed entities will be significant. These adsorbed entities form the surface states, and then rate control at the surface (not transport in the interior) becomes rate determining for the overall photoelectrochemical reaction. Before going further with the question of the identity and the location of rate control in photoelectrochemistry, it is better to discuss what surface states are and how one finds out about them.

10.5.2. Surface States

10.5.2.1. *Introduction.*   So far, references in this chapter to surface states have been anticipatory; now we are going to describe them. But first, it is helpful to present an analogy.

From the structure of the interfacial region presented in Chapter 6, it follows that the second approximation of its structure, the 1916 Gouy theory, consists in regarding the electric charge carried by the solution component of the interfacial region as distributed according to Boltzmann's law in a "charge cloud," (the Gouy diffuse layer) while the charge on the electrode remains adherent to a single layer on the surface of the electrode (Section 6.7.4). Later on, it was realized that this picture (though giving rise to elegant mathematics) did not present the heart of the matter, but rather that this was usually more to be found in the Helmholtz layer, suggested earlier i.e., that the interfacial structure is just a "double layer," the excess electric charge on the metal being balanced by a single-layer counter-charge on the solution side. There are circumstances (concentrated solutions) where this Helmholtz layer dominates the interface, and there are circumstances (extremely dilute solutions) in which the diffuse Gouy part dominates.

However, the reason for this look back at some of the contents of Chapter 6 is the nature of ions adsorbed at metals. They are of two types. The Gouy diffuse-layer ions simply exist in the solution bulk, but lack any specific chemistry, being either positive or negative, depending on the charge on the electrode and in the first (or Helmholtz) layer. In this latter region, some of the ions turn out to be bound *chemically* (contact adsorption) to the surface of the electrode itself.

In the semiconductor/solution interface, the situation is analogous to that just reviewed. In very dilute solutions and particularly when only redox ions are present (no H_3O^+ or OH^- dominating the interface), the electrons and holes on the semiconductor side are in fact distributed in a charge cloud within the semiconductor (Fig. 10.2); compare the ions in the charge cloud of the Gouy theory. However, if the solution becomes more concentrated, but particularly if H or O can be adsorbed on the semiconductor surface, the charge distribution within the electrode may change. It is no longer all distributed in an exponential manner with respect to distance, like the distribution of the positive and negative ions in the diffuse layer on the solution side. In fact, the surface states at the interface provide an analogue of the specifically adsorbed ions in the inner double layer (Section 6.7.5). Thus, the distribution of potential (and hence charge) at the semiconductor solution shown earlier (Fig. 10.1) is a valid presentation of the situation in a semiconductor without surface states. However, in aqueous solutions and in practical concentration ranges, surface states on semiconductor electrodes often predominate.

Surface states, then, are states in or near the surface of semiconductors that can be occupied by electrons or holes. They are always discussed as existing energetically *in* the energy gap; they are most noticeable here because in the first approximation of the Schottky model, there should be no states for electrons at all in this energy region.

Surface effects cause important changes in the model we have for photoelectro-chemical reactions. Thus, when the surface state concentration exceeds 10^{12} cm^{-2}, it begins to change the structure of the semiconductor/solution interface. Sucking the charge out of the space charge region to put it onto the semiconductor surface changes the shape of the charge distribution inside the semiconductor and diminishes the importance of the space charge region. In the presence of a sufficiently high concentration of surface states (most of the charge now on the surface), the semiconductor/solution interface begins to resemble the metal/solution interface, and the potential difference—which in the Schottky model was within the semiconductor—is moved out into the solution. In turn, the location of the largest part of the potential difference at the interface will be Helmholtz-like and the drop within the semiconductor will be relatively small (Fig. 10.15).

10.5.3. Determination of Surface States

There are several approaches to determining surface states. The first depends upon the violation, owing to the presence of surface states, of the first rule of photoelectro-chemistry: the photons in the irradiating light must have energies big enough to knock
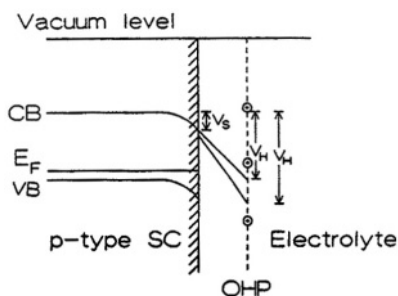
Fig. 10.15. *p*-Type semiconductor/solution interface in the presence of high-density surface states. The potential difference in the Helmholtz part of the double layer, (i.e., that in the solution) is greatly increased compared with a situation with a negligible number of surface states. Correspondingly, the potential difference within the semiconductor is greatly diminished compared with one containing negligible surface states.

electrons up out of the valence band and into the conduction band. Insufficiently energetic photons would simply fall back into the valence band.

However, in a second approximation one admits that there may be some extra states on the surface that are energetically within the band gap. Then irradiation with light that has photons of *less* than the band gap energy may also cause a photocurrent to flow. Thus, if a surface state is in the middle of the energy gap, E_g, the energy needed to bring up an electron to it is not E_g, but $1/2\ E_g$. Therefore, if one carefully determines the minimum frequency of light needed in a monochromatic irradiating beam to produce a photocurrent, one can assess the energy of the surface state with respect to that of the valence band. The intensity of the photocurrent density at that energy may help us obtain the concentration of the surface state.

Another approach involves impedance analysis (Section 7.5.13). One measures impedance as a function of frequency of the applied current and finds that (for the imaginary impedance, say) there is an unexpected maximum on the Z'–log ω plot. Analysis of the data allows one to numerically isolate the unexpected "anomaly" in the impedance plot, obtain the equivalent capacitance and resistance, and then interpret these in a model as representing a surface state.

Third, one may use a scanning-tunneling approach (Section 7) and record the *i–V* plot. For example, in one such analysis using *p*-Si in an alkaline solution, a state was found at 0.3 V below the energy level of the conduction band (Fan and Bard, 1990).

A somewhat different use of scanning tunneling microscopy to measure surface states has been made by Szklarczyk and Gonzalez-Martin (1991) and by Uosaki (1995) (see Figs. 10.16 and 10.17).
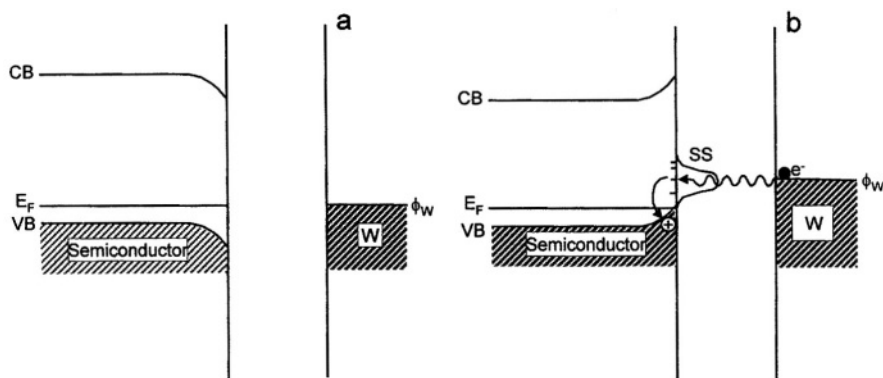
Fig. 10.16. Energy band diagrams for: (a) Schottky barrier situation and (b) tunneling situation. *CB*, conduction-band energy; *VB*, valence-band energy; E_F, Fermi energy level. (Reprinted from A. Gonzalez-Martin, thesis, Texas A&M University, 1993.)
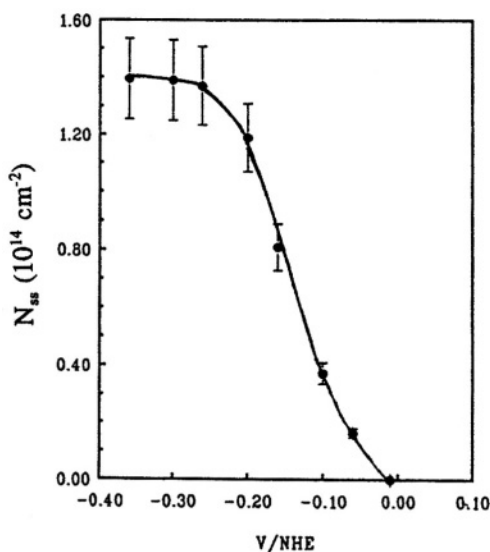
Fig. 10.17. N_{ss} as a function of the electrode potential. CdTe as electrode; 5% aqueous dimethyl formamide (DMF) Tetraalkylammonium perchlorates present. (Reprinted with permission from J. O'M. Bockris and J. Wass, *Mater. Chem. Phys.* **22**: 249, 1989).

10.5.4. What Causes a Surface State?

Surface states on a semiconductor in a vacuum can sometimes be explained by means of the spare bonds that "dangle" from atoms on surfaces, or defects associated with dislocations. Neither of these mechanisms works at the semiconductor/solution interface. The dangling bonds will be expunged by adsorbed water, etc. Experiment shows that the concentration of surface states on semiconductors in solution is strongly potential dependent, and that defects in the crystal structure would not be potential dependent, at least until anodic dissolution of the substrate itself began.

Ion adsorption has been suggested as a cause of surface states (Rajeschwar, 1982) and it may well be one cause, particularly if the potential of the semiconductor is positive to that of the flatband potential. On p-Si in a solution containing $N(Bu)_4^+$, to take a specific example, the density of surface states obtained from impedance spectroscopy is shown in Fig. 10.17 (Gonzalez-Martin, 1993). It is seen that the concentration of surface states increases exponentially in the cathodic direction until the expected limit (toward complete coverage) is seen.

For equilibrium in the probable proton discharge onto a thin (5 Å) layer of oxide on Si,

$$A^+ + e \rightleftarrows A_{ads} \tag{10.3}$$

$$k_1[A^+](1 - 0_A) \exp\left(-\frac{\beta\eta F}{RT}\right) = k_{-1}\theta_A \exp\left(\frac{(1-\beta)\eta F}{RT}\right) \tag{10.4}$$

and

$$\theta_A = \frac{1}{1 + \dfrac{1}{K[A^+]} \exp\left(\dfrac{F\eta}{RT}\right)} \tag{10.5}$$

where θ_A is the fraction of the electrode covered with adsorbed A in equilibrium (A can be the H atom), k_1 and k_{-1} are the rate constants, and K is an equilibrium constant for the reaction in Eq. (10.3) ($K = k_1/k_{-1}$). At $\eta = 0$, from Eq. (10.3):

$$\frac{\theta_{\eta=0}}{1 - \theta_{\eta=0}} = K[A^+] \tag{10.6}$$

For a nonpolarizable double layer, the charges led in or out of the electrode from the outer circuit are equivalent to a capacitance $C = \partial q_H/\partial V$). However, this would be a pseudo-capacitance because the charge penetrates the double layer. C_{ads} can be calculated from Eq. (10.5).

$$C_{ads} = \frac{\partial q_H}{\partial\eta} = -\frac{\partial}{\partial\eta}FZ\theta_{Hads} = \frac{F^2 Z}{RTK[A]} \frac{\exp\dfrac{F\eta}{RT}}{\left[1 + \dfrac{1}{K[A^+]} \exp\left(\dfrac{F\eta}{RT}\right)\right]^2} \tag{10.7}$$

where Z is the number of sites on the surface ($\approx 10^{15}$ cm^{-2}). The maximum value of C_{ads}, $C_{ads,max}$ can be estimated from the relation $\partial C_{ads}/\partial\eta = 0$ [Eq. (10.7)].

At $\dfrac{\partial C_{ads}}{\partial\eta} = 0$,

$$C_{ads,max} = \frac{F^2 Z}{4RT} = 1600\mu F \text{ cm}^{-2} \tag{10.8}$$

An estimate of the actual coverage can be obtained by comparing the maximum value of C_{ads} (i.e., by the model) with C_{SS} (i.e., experimental, 465 $\mu F\,cm^{-2}$). $C_{ads,max} = 1600\ \mu F\,cm^{-2}$ is calculated using Eq. (10.8) (taking $Z = 1 \times 10^{15}\ cm^{-2}$). $C_{SS,max} = 465\ \mu F\,cm^{-2}$. Thus,

$$\theta = \frac{465}{1600} = 0.3 \qquad (10.9)$$

Surface states may lead to a reduction of a photocurrent if the state leads to the trapping and deactivation of electrons or holes (which would otherwise have reached the semiconductor solution interface ("bad" surface states). On the other hand, if surface states in the band gap allow charge transfer to or from surface states in the solution that would not have been otherwise possible, the photocurrent will be increased ("good" surface states). A possible mechanism for a bad surface state is shown in Fig. 10.18. A good surface state mechanism is shown in Fig. 10.19.

10.5.5. The Effect of Surface States on the Distribution of Potential in the Semiconductor Interface

It is possible to determine how surface states affect the distribution of potential in the interface if one assumes that the solution concerned is relatively concentrated so that the excess electric charge on the solution side of the interface is predominantly in the Helmholtz (H) layer. Then

$$q_{sc} = q_H$$

One assumes here that the electrolyte concentration is high and that the potential drop in the diffuse layer is negligible compared with that in the Helmholtz layer.
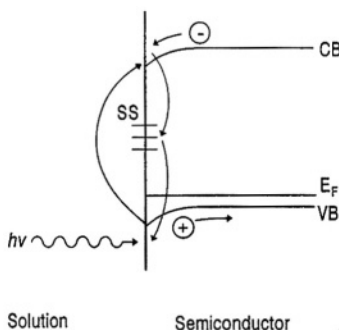
Fig. 10.18. Representation of a "bad" surface state on a semiconductor electrode. (Reprinted from A. Gonzalez-Martin, thesis, Texas A&M University, 1993.)
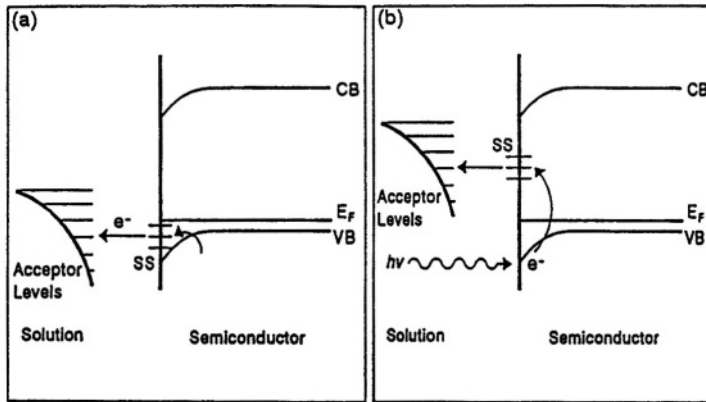
Fig. 10.19. Representation of a "good" surface state on a semiconductor electrode. (a) Thermal activation and (b) photoexcitation of an electron from the valence band to surface states. (Reprinted from A. Gonzalez-Martin, thesis, Texas A&M University, 1993.)

The charge in the space charge (sc) layer can be calculated by solving Poisson's equation:

$$d_{v\,dx}^2{}^2 = -(4\pi\varepsilon)p_x \tag{10.10}$$

The charge density, p_x, is made up of N_{sc} positive charges in the depletion layer and $n = N_{sc}e^{-e_0 VkT}$ electrons in the conduction band. Therefore,

$$\frac{d^2V}{dx^2} = -\frac{4\pi}{\varepsilon}e_0N_{sc}(e^{-e_0V/kT} - 1) \tag{10.11}$$

However,

$$\frac{d^2V}{dx^2} = \frac{1}{2}\frac{d}{dV}\left(\frac{dV}{dx}\right)^2 \tag{10.12}$$

Integration of Eq. (10.11) gives

$$\left(\frac{dV}{dx}\right)^2 = -\frac{8\pi}{\varepsilon}e_0N_x(-e^{-e_0V/kT}(kT/e_0) - V) + \text{constant} \tag{10.13}$$

But $dV/dx = 0$ when $V = 0$, giving

$$\text{const} = -(8\pi/\varepsilon)e_0N_{sc}(kT/e_0) \tag{10.14}$$

At the semiconductor surface, $V = V_{sc}$ and the factor $e^{-e_0 V_{sc}/kT}$ is negligible. Thus,

$$(dV/dx)_{V=V_{sc}} = [(8\pi/\varepsilon)e_0 N_{sc}]^{1/2}[V_{sc} - (kT/e_0)]^{1/2} \qquad (10.15)$$

From Gauss's law,

$$(dV/dx)_{V=V_{sc}} = (4\pi/\varepsilon)q_{sc} \qquad (10.16)$$

Therefore,

$$q_{sc} = [\varepsilon e_0 N_{sc}/(2\pi)]^{1/2}[V_{sc} - (kT/e_0)]^{1/2} \qquad (10.17)$$

Also:

$$V_H = (q_H/C_H) = [q_{sc}(4\pi d_H/\varepsilon_w)] \qquad (10.18)$$

Taking $\varepsilon_w = 6$, $\varepsilon = 10$, $d_H = 3$ Å and $N_{sc} = 10^{16}$ cm^{-3}, $\qquad (10.19)$

$$V_H = 1.73 \times 10^{-4}[V_{sc} - (kT/e_0)]^{1/2} \qquad (10.20)$$

Thus the numerical value of V_H (the order of magnitude unity at a metal/solution interface) is here reduced to the order of tens of millivolts, which is negligible compared with the potential difference inside the semiconductor.

In the presence of sufficient surface states, the charge on the solid side of the interface is distributed in the space charge region and the surface states. Thus,

$$q_H q_{sc} + q_H \qquad (10.21)$$

The charge in the surface states is given by

$$q_H = \varepsilon_0 N_{SS} \qquad (10.22)$$

where N_{SS} is the number of surface states per unit area. The potential drop in the Helmholtz region in the presence of surface states then becomes

$$V_H = (q_{SS} + q_H)(4\pi d_H/\varepsilon_w) \qquad (10.23)$$

The relative changes in V_H and V_{sc} as a function of surface state density are shown in Fig. 10.20. At low surface state density ($<10^{12}$), the potential drop across the Helmholtz layer is small and remains almost constant with a change in electrode potential. However, at high surface state densities ($>10^{13}$), the potential drop in the Helmholtz region increases and exceeds the potential drop in the space charge region for surface state densities greater than 5×10^{13} cm^{-2}.

At sufficiently high doping concentrations ($\sim 10^{19}$ cm^{-3}) or sufficiently high density of surface states, the ratio $V_{Helmholtz}/V_{semiconductor}$ exceeds unity, and the prop-
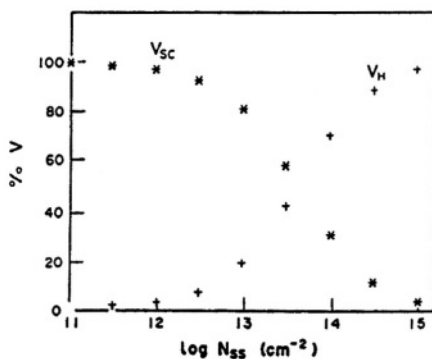
Fig. 10.20. Relative potential drop in the space charge region and in the Helmholtz region as a function of surface state density. (Reprinted from K. Chandresakaran, R. C. Kainthla, and J. O'M. Bockris, *Electrochim. Acta* **33**: 334, Fig. 12, copyright 1988, with permission from Elsevier Science.)

erties of the semiconductor/solution interface become similar to those of the metal/solution interface.

10.5.6. Kinetic Photoelectrochemical Processes at High Surface State Semiconductors

There is a similarity in the course of developments in photoelectrochemical kinetics since the mid-1980s and that of the gradual taking into account of adsorbed intermediate radicals in thermal electrode kinetic theory during the last two decades of the century. Thus, outer sphere reactions (influenced only by a force *outside* the first layer of water molecules) are in practice rare, although most of the development of the quantum theory of electrode kinetics (D. Miller, 1995; Schmickler, 1996) is based upon them. Similarly, photoelectrochemical reactions that are controlled by diffusion inside the electrode (the Schottky case) *outside* their limiting current region also seem rare. Indeed, encouraging them to occur may require nonaqueous solutions because the most frequent cause of the plentiful surface states that require the Helmholtz model[12] for charge distribution at the interface is adsorbed H, which is

---

[12] Jaegerman (1997) calls photoelectrochemical reactions that occur at high surface state configurations "Bardeen model reactions" because Bardeen was the first to discuss surface states. Here they are called *Helmholtz reactions* because the potential difference at interfaces at which they predominate is largely in the Helmholtz layer.

absent in anhydrous solutions. Thus, just as most electrochemical reactions involve intermediate radicals adsorbed on the electrode, so most photoelectrochemical reactions involve a high surface state concentration. This was first described in kinetic equations by Bockris and Uosaki (1977).

An outline of Butler's theory for the terms of the low surface state (transport-controlled) case is given in Section 10.3.5. Uosaki's 1977 theory of kinetics in the high surface state case was developed in greater detail by Khan (1984). Here, the beginning equation for the steady state ($\partial n_x/\partial t = 0$) in the space charge region has a flux independent of distance, so that from the Nernst–Planck equation (4.226) and with $dJ/dx = 0$, one obtains:

$$\frac{d^2 n_1(x)}{dx^2} - \frac{e_0}{kT} V'(x) \frac{dn_1(x)}{dx} - \frac{e_0}{kT} V''(x) n_1 x = \frac{(1 - R_\nu) I_\nu \alpha_\nu \exp(-\alpha x)}{D_e} \quad (10.24)$$

Here, $n_1(x)$ is the carrier concentration at a distance x from the surface, R is the coefficient of reflection, I_ν is the intensity of illumination at a frequency ν and α_x is the absorption coefficient for photons, while D_e is the diffusion coefficient of electrons. This equation can be understood if one recalls Fick's second law, which gives the rate of change in concentration during diffusion in the absence of an electric field (Vol. 1, Section 4.2). The other terms on the left represent the application of the Nernst–Planck equation (Section 4.4). The term on the right represents the rate of absorption of light, taking into account a reflection coefficient, R_ν at a frequency ν.

One also has to consider charge generation in the field-free region for which

$$\frac{d^2 n_2}{dx^2} - \frac{n_2(x)}{L_D^2} = -\frac{(1 - R_\nu) I_\nu \alpha_\nu e^{-\alpha_\nu x}}{D_e} \quad (10.25)$$

because here there is no electric field and hence $V' = V'' = 0$.

Solution of these equations eventually gives an expression for $n_1(\theta)$, the concentration of electrons at the surface, in terms of the parameters already defined, together with k_{CT}, the rate constant for charge transfer across the interface (see Fig. 10.21) and two recombination constants, one for the bulk and one for the surface. Recombination of hole–electron pairs is taken into account in the development, as is also the formation of surface states by a surface-dependent anion adsorption at a degree of coverage, θ.

The charge-transfer rate constant is expressed in Gurneyan terms (Sec. 9.6) by the equation:

$$k_{CT} = S_e \int_{E_c}^{\infty} P(E) f(E, h\nu) D(E_0, E) dE \quad (10.26)$$

where S_e is the drift velocity of the outgoing electrons in the surface region of the semiconductor and $N(E_0, E)$ is the density of acceptor states in solution.
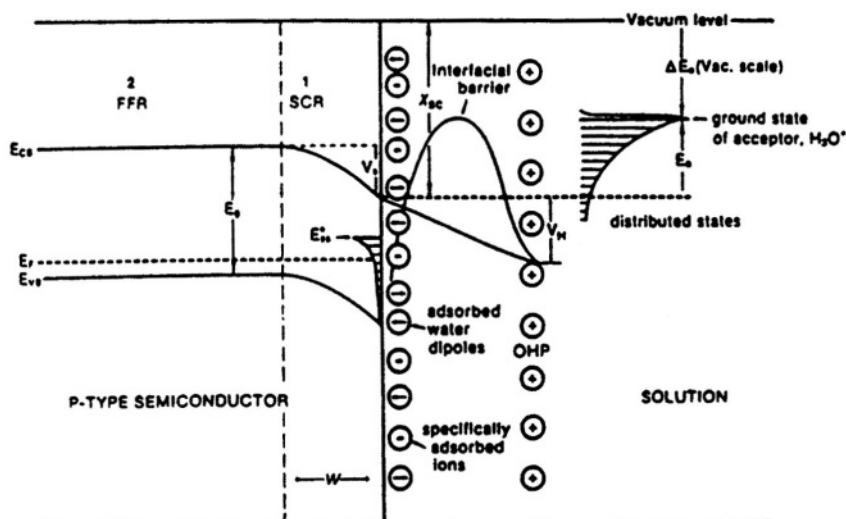
Fig. 10.21. A schematic representation of a *p*-type semiconductor/solution interface. (Reprinted from S. U. M. Khan and J. O'M. Bockris, *Int. J. Hydrogen Energy*, **11**: 373, 1986. Reproduced with permission of T. Nejat Veziroglu.)

One finally arrives at a somewhat cumbersome expression for the photocurrent:

$$i_{\text{ph}}(\nu) = e_0 k_{\text{CT}} n_1(0) = \frac{e_0 k_{\text{CT}}(1 - R_\nu) I_\nu}{k_{\text{CT}} + k_{\text{sr}} + k_{\text{br}}} \times \left[1 - \frac{\exp(-\alpha_\nu W)}{(1 + \alpha_\nu L_D^{-1})[1 + L_D^{-1} G_1(W)]} - \frac{G_2(W)}{L_D + G_1(W)} \right]$$

(10.27)

where W represents the width of the space charge region and G_W is a parameter depending on W and ΔV_{SC}, the remaining (small) potential difference inside the semiconductor, in the presence of surface states; k_{sr} and k_{br} are the rate constants for the recombination of electrons and holes in the surface and bulk, respectively.

Equation (10.27) indicates that the charge transfer becomes the rate-limiting step under the condition when $k_{\text{CT}} \ll (k_{\text{sr}} + k_{\text{br}})$. The term in large brackets is a function of transport control of the photocurrent. If the electrode potential is sufficiently negative in a cathodic reaction at a *p*-type semiconductor, $k_{\text{CT}} \gg (k_{\text{sr}} + k_{\text{br}})$ and interfacial charge transfer control is lost. Eventually, control passes to transport within the semiconductor (although it is affected by recombination).

This expression has to be finally integrated over the region in which photoexcitation begins, $E_g/h = \nu_{\text{crit}}$. Therefore,

$$i_{ph}(\text{solar}) = \int_{\nu_{\text{crit}}}^{\nu_{\text{max}}} i_{ph}(\nu) d\nu$$

(10.28)

This expression has been evaluated (Khan, 1984) and gives the S-shaped curve of experiment (Fig. 10.22).

10.5.7. Looking Back and Looking Forward at Photoelectrochemistry

We have already had a review (Section 10.3.5) of the first approximation for the photoelectrochemistry that occurs when semiconductors in solution are irradiated—that in which the photocurrent–potential relation is considered to be near the limiting current region, so that rate control lies in the transport of carriers from the interior of the semiconductor where they are produced upon absorption of photons, to the interface. We learned of the importance of a match between the energy gap E_g of semiconductors and the maximum of the solar spectrum (a fair match occurs for abundant silicon). In the poorly named Schottky barrier model, the photons of the incoming light that manage to get electrons up from the valence band into the conduction band produce electrons as minority carriers (e.g., in p -type cathodes), and the rate-determining step for the photocurrent is the transport of these carriers to the surface of the semiconductor. Such a model was developed in mathematical form by Butler in 1977, but is limited in applicability because it neglects surface recombination and surface states.

Since 1980, the influence of surface properties of semiconductors on photoelec-trochemical properties has been taken into account, and this has led to investigations of surface states (Chandresakaran, 1985; Chazavil and Rajeschwar, 1990s; Hamnett,
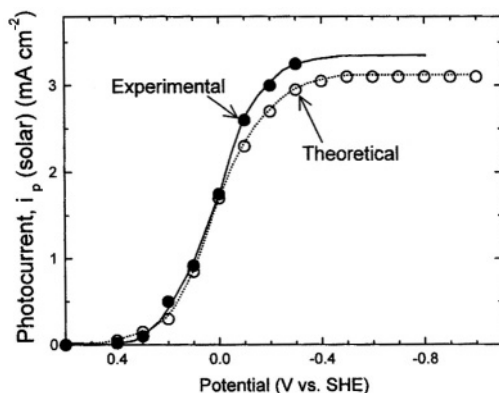
Fig. 10.22. Plot of photocurrent, i_{ph} (solar), as a function of electrode potential, V (volts) vs. NHE. (Reprinted with permission from S. U. M. Khan and J. O'M. Bockris, *J. Phys. Chem.* **88**: 2511, Fig. 3, Copyright 1984, American Chemical Society.)

1992). Known earlier from investigations of the semiconductor surface/gas interface, surface states are caused at the semiconductor/solution interface by intermediate radical formation (particularly of H and O), but also by the contact adsorption of ions. When there is a sufficient concentration on the surface ($\theta > 0.01$), the structure of the interfacial region begins to change and becomes eventually for $\theta > 0.1$ like that of a metal, with much of the potential difference around the surface now on the solution side rather than inside the semiconductor, as is the case with the Schottky model. Now the rate-determining step tends to change and (except in the limiting current region) becomes charge transfer at the semiconductor/solution boundary. Such a conclusion is emphasized by the observation (Uosaki and Kita, 1982) that the photoelectrochemical current below the limiting region obeys Tafel's law. Such a situation must take into account the recombination of hole–electron pairs inside the semiconductor[13] and the potential dependence of the coverage of the surface with surface states. Such a scenario prepares us for an early look at the usefulness of semiconductor photoelectrochemistry in the fixing of CO_2, the splitting of water by solar light, and the elimination of toxic wastes, etc.

10.6. PHOTOELECTROCATALYSIS

Photoelectrocatalysis refers to the accelerative effect on photoelectrochemical reactions sometimes achieved by decorating part of the semiconductor surface with small islets of metals (Tsubmura, 1976). The effect is illustrated for the photoevolution of O_2 on TiO_2 decorated with a number of metals. In Fig. 10.13 it was shown that this effect works for O_2 evolution in alkaline solutions, but there is no corresponding effect in acid solutions (Contractor and Bockris, 1987). An early attempt to account for these effects on the basis that the metal islets changed the Fermi level of the semiconducting substrate failed when it predicted that catalysis would be best with islets of metals having low work functions (Φ); in fact, metals of high Φ work best.

The following evidence suggests a purely catalytic interpretation of the effect of metal islands on the velocity of hydrogen and oxygen evolution, at least in some cases. The model assumes that a high concentration of surface states is present. It therefore follows that the potential difference in the interphase (e.g., for p-Si decorated with Pt islets evolving H_2) is larger in the Helmholtz layer and smaller inside the semiconductor. The photoproduced electrons are transported to the surface by relatively fast transport processes, the minority carriers being in fact in a kind of pseudo-equilibrium with those in the bulk. At the semiconductor/solution interface, electron transfer (e.g., to H_3O^+) occurs. Electrons find this (rate-determining) step easier at the metal catalyst/solution boundary than at the semiconductor/solution interface. The kinetics

---

[13]The lessened value of dV/dx within the semiconductor in the high surface state case tends to increase the loss of active electrons by means of hole–electron pair recombination because the rate of spatial separation of the electron from its hole, which was formerly encouraged by the electric fields sending the oppositely charged entities in the opposite directions, is now diminished.
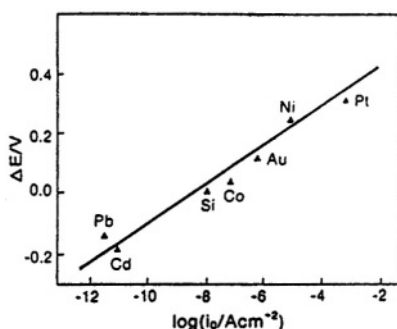
Fig. 10.23. The dependence of the ΔE shift caused by the presence of metal islets on log i_0 for hydrogen evolution in the dark on the corresponding massive metals. (Reprinted from M. Szklarczyk and J. O'M. Bockris, *Appl. Phys. Lett.* **42**: 1042, 1983.)
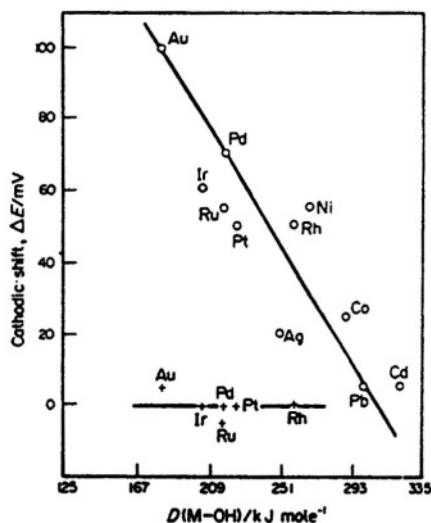
Fig. 10.24. Dependence of the cathodic displacement ΔE on the bond energy $D(\text{M–OH})$. (Reprinted from A. Q. Contractor and J. O'M. Bockris, *Electrochim. Acta* **32**: 121, Copyright 1987, with permission from Elsevier Science.)

of the photoevolution reaction therefore become characteristic of thermal electrode reactions at interfaces of solutions with the islet metals concerned. The semiconductor simply acts to supply charge carriers as a result of its absorption of light.

Such a catalytic interpretation of these effects of metal islets is consistent with the fact that the lessening of overpotential for hydrogen or oxygen at a given rate increases linearly with $\log i_0$ where i_0 is the exchange current density for the corresponding thermal reaction (Fig. 10.23). It is consistent also with the lack of catalytic effects for oxygen evolution in acid solution. Thus, rate-determining discharge of water onto Ti atoms in the oxide surface would be faster than onto any of the noble metal islets added to the surface because the $Ti{=}O$ bond is stronger than any noble metal–O bond and in acid solution on Ti, it is discharge that is the rate-determining step in oxygen evolution. However, in alkaline solution, the rate-determining step for oxygen evolution involves *desorption* from O bound to M. The metals added to the TiO_2 now offer a faster path for O_2 evolution because desorption of O from the weaker M–O bonds is easier than from the TiO bonds. Accordingly, an increasingly fast reaction occurs as the M–O bond decreases in strength (see Fig. 10.24). In general, the effects of decoration of semiconductor surface catalysts seem to occur as shown in as Fig. 10.25 for the photoreduction of CO_2.
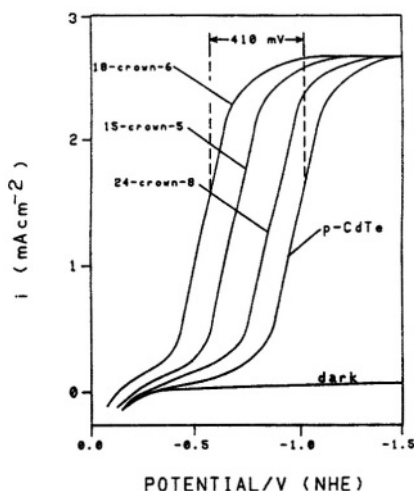
Fig. 10.25. Current–potential curves for crown-ether catalysts added to the electrolyte for a *p*-CdTe electrode in DMF–0.1 *M* TEAP(tetraethylammonium phosphate)/5% H_2O. (Reprinted from J. O'M. Bockris and J. W. Wass, *J. Electrochem. Soc.* **136**: 2523, 1989. Reproduced by permission of the Electrochemical Society, Inc.)

Photoelectrocatalysis *transforms the possibilities of using photoelectrochemistry in practical situations.* For example, it increases the efficiency of conversion of light to hydrogen by more than one order of magnitude and brings it to values competitive with those of photovoltaic conversions of light to electricity (see Section 10.7.1.).

10.7. THE PHOTOELECTROCHEMICAL SPLITTING OF WATER

10.7.1. The Need for Photoelectrocatalysis

In order to obtain pure hydrogen for use in fuel cells in transportation and for other purposes, it would be environmentally advantageous to obtain the hydrogen from abundant water, because then the co-product is oxygen and no CO_2 is evolved into the atmosphere. Starting in 2003, Daimler-Benz expects to mass produce more than 100,000 electric cars per year fueled by methanol or gasoline, with on-board re-forming to H_2 for use in the fuel cells producing electricity for the motors. However, such a system of on-board re-forming will still emit CO_2 to the atmosphere from the re-forming reaction, although it will be about half that now emitted per passenger mile, owing to the roughly doubling of efficiency in energy conversion of the electrochemical engine (the fuel cell–electric motor combination) over that of a Carnot-limited heat engine, the internal combustion motor.

In order to end the current accelerated trend toward planetary warming, CO_2 emissions into the atmosphere from land and air transport must be eliminated, and the way to do this is simple in principle: Make hydrogen from water by photo-splitting water using solar light (or electricity from a nuclear source).

As explained earlier, photoelectrochemical splitting of water was done for the first time in 1972 (Fujishima and Honda). However, the efficiency of this cell was very low (about 1%) and hence not practical. A number of advances have brought an economical standalone, one-step solar water-splitting technology much nearer. There have been four steps in these advances.

The first step was the evolution away from the Schottky barrier model of photoelectrochemistry caused by the evidence from the late 1970s onward that the rate of photoelectrochemical reactions was heavily dependent on surface effects (Uosaki, 1981; Szklarczyk, 1983). This was followed by the use of both a photocathode and a photoanode in the same cell (Ohashi, 1977). Then the use of nonactive thin protective passive layers of oxides and sulfides allowed photoanodes to operate in potential regions in which they would otherwise have dissolved (Bockris and Uosaki, 1977). The final step was the introduction of electrocatalysis of both hydrogen and oxygen evolution by means of metal islets of appropriate catalytic power (Bockris and Szklarczyk, 1983).

It was nevertheless necessary to introduce a theory for electrode matching because the work by Uosaki and Ohashi had shown (individually) that the efficiency of the one

electrode in a cell was strongly dependent on the characteristics of the co-irradiated counter-electrode. Such a theory was not developed until 1986 (Khan and Kainthla).

The first standalone photo water-splitting device having a practical conversion efficiency was described by Kainthla and Khan (1987). Their cell is shown in Fig. 10.26. Both electrodes were simultaneously irradiated. The calculated optimal match between materials of the two electrodes was found in theory to be p-InP as cathode and n-GaAs, the latter protected from electrochemical oxidation by a film of MnO_2. The InP was duly decorated with platinum to catalyze H recombination; the MnO_2 allowed photo O_2 evolution on its surface with a stability unchanged over prolonged periods.

A plot of the efficiency of conversion of solar light to clean hydrogen fuel from water is shown in Fig. 10.27. A maximum efficiency of 8.4% was obtained, and this could be increased to 9.3% by increasing the temperature to 43°C. In 1998 Turner et al. improved upon this and obtained 16% efficiency in the catalyzed photoelectro-chemical splitting of water.

Photoelectrochemical standalone one-step photo water-splitting technology may be finally economically superior to the photovoltaic production of electricity followed by the electrolysis of water. The advantage of the photoelectrochemical pathway to massive photoproduction of clean hydrogen as a fuel lies in the superior potential economics of the single plant (compared with the two plants needed in using photo-voltaics to electrolyze water). There would be a further gain in the elimination of the efficiency losses that arise from having to multiply the efficiency of the photovoltaics (10–20%) by that of the electrolyzer (70–80%). Thus, using the middle ranges of these
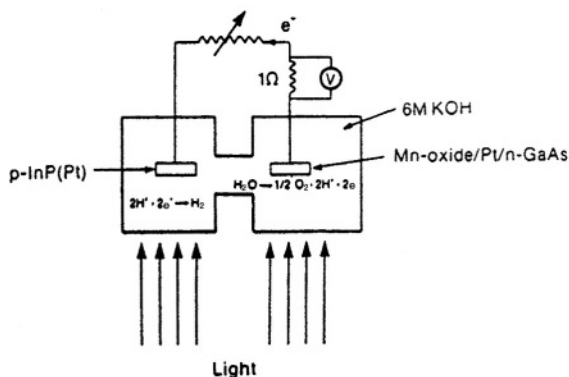
Fig. 10.26. Schematic diagram of the self-driven pho-toelectrochemical cell for water electrolysis. (Re-printed from R. C. Kainthla, S. U. M. Khan, and J. O'M. Bockris, *Int. J. Hydrogen Energy* **12**: 381, 1987. Re-produced with permission of the International Associa-tion for Hydrogen Energy.)
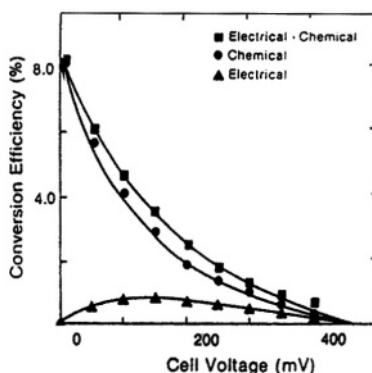
Fig. 10.27. Chemical, electrical, and total (chemical+electrical) efficiency of self-driven phtoelectrochemical cell employing p-InP (Pt) and MnO_2-coated n-GaAs electrodes in 6 M KOH solution as a function of cell voltage. (Reprinted from R. C. Kainthla, S. U. M. Khan, and J. O'M. Bockris, *Int. J. Hydrogen Energy* **12**: 381, 1987. Reproduced with permission of the International Association for Hydrogen Energy.)

figures, 15 and 75%, respectively, 10.5% would be the net result from the dual approach. This figure is already less than laboratory values obtained in photoelectrochemical water splitting.

To produce clean hydrogen fuel by photosplitting water on a massive scale, sea-based solar collection of hydrogen would be desirable. Such possibilities have been evaluated by Ichikawa (1994). It should be possible to avoid the evolution of Cl_2 ($E_0 = 1.35$) gas into the atmosphere by using brine (highly concentrated sea water) and then controlling the anode potential to values less anodic than 1.4 V. This would make the principal anodic reaction the formation of liquid bromine, which would then eliminate the danger of Cl_2 reaching the atmosphere.

10.7.2. Could Cheap TiO_2 Be Used in the Economic Photoelectrolysis of Water?

Economic considerations will play the main part in deciding on the methods to be used in eliminating the use of oil and coal as energy sources. In discussing

photoelectrochemical devices in Section 10.7.1, it was maintained that a maximum efficiency of 16% solar light conversion to H_2 has been reached. The photovoltaic-electrolyzer combination would be unlikely to have an efficiency greater than this.[14] Further development of the two photo devices using photoelectrocatalysis may produce efficiencies exceeding that of a photovoltaic–electrolyzer combination if, for example, Ni-decorated Si semiconductors were used in the photocathode. However, the use of a protected GaAs anode as a photoanode is still an obstacle because of the high cost of GaAs.

The cheapest photoelectrode material by far is TiO_2. It is plentifully available from the mineral ilmenite, K_2TiO_3. The disadvantage of TiO_2 is its high E_g value of ~3 eV (see Fig. 10.11). Comparison of this value with the solar spectrum shows that when TiO_2 is irradiated with solar light, it absorbs only about 2% of this light. It would seem to be unpromising as a component of a potential commercial water splitter.

M. Grätzel (1992), however, has taken the attitude that the economic attractiveness of TiO_2 is so great compared with that of other semiconductors that it might pay to try to modify TiO_2 until it does the necessary job. Grätzel and his research team have made progress in two directions:

1. They have produced *faceted* TiO_2 (an increase in surface area of about 1000 times has been achieved). When light strikes a semiconductor, it is reflected as well as absorbed. If the semiconductor surface concerned is flat, the reflected light is lost, but if the surface is made rough in a suitable way, the reflected light strikes the semiconductor surface more than once, with a consequent increase in absorption (Fig. 10.28).

2. It has long been a concept that one might be able to increase the solar absorption efficiency of a semiconductor by decorating the surface with a dye having suitable photoreceptors in wavelengths outside those in which the semiconductor adsorbs. Certain thiocyanates of organics involving bipyridyl complexed to ruthenium perform best (Fig. 10.29). A concern in an electrochemical anode (TiO_2 is an *n*-type semiconductor) operating in aqueous solutions is the probable eventual anodic oxidation of the dye (Grätzel, 1993).

It has at least been shown by Grätzel and co-workers that significant (i.e., more than 10 times) increases in the efficiency of TiO_2 as a photoanode can be obtained using these two modifications, though in nonaqueous solutions. Such devices may therefore yet present competition to the two photon electrocatalyzed anode-protected devices that have shown such high efficiency and that use aqueous solutions. It is a matter of the efficiency obtained versus the cost of the materials—and of the necessary plant.

---

[14]Of course, this depends on the assumptions made about the efficiency of photovoltaic energy conversion. The efficiency of an electrolyte is ~0.7. If photovoltaic efficiencies of 0.3 or greater could be reached, a 20% overall efficiency of hydrogen from this route would be possible.
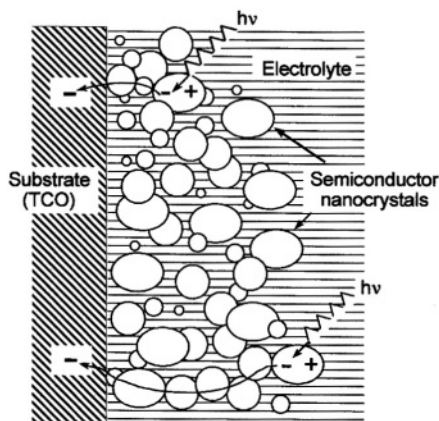
Fig. 10.28. Model of charge carrier separation and charge transport in a nanocrystalline film. The electrolyte has contact with the individual nanocrystallites. Illumination produces an electron–hole pair in one crystallite. The hole transfers to the electrolyte and the electron traverses several crystallites before reaching the substrate. Note that the photogenerated hole always has a short distance (about the radius of the particle) to pass before reaching the semiconductor/electrolyte interface wherever the electron–hole pair is created in the nanoporous film. The probability for the electron to recombine will, however, depend on the distance between the photoexcited particle and the tin-coated oxide back-contact. (Reprinted with permission from A. Hagfeldt and Michael Grätzel, "Light-Induced Redox Reactions in Nanocrystalline Systems" *Chem. Rev.* **95**: 49–68, copyright 1995, American Chemical Society.)
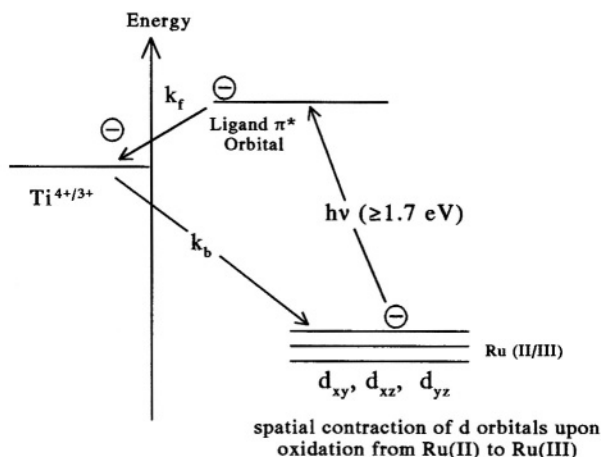
Fig. 10.29. Molecular orbital diagram for ruthenium com-
plexes anchored to a TiO_2 surface by a carboxylated
bipyridyl ligand. The visible light absorption of these types
of complexes is a metal-to-ligand charge transfer (MLCT)
transition. The carboxylate groups are directly coordi-
nated with the surface titanium ions, producing intimate
electronic contact between the sensitizer and the semi-
conductor. (Reprinted with permission from A. Hagfeldt
and Michael Grätzel, "Light-Induced Redox Reactions in
Nanocrystalline Systems," Chem. Rev. 95: 49–68, copy-
right 1995, American Chemical Society.)

10.8. THE PHOTOELECTROCHEMICAL REDUCTION OF CO_2

The reduction of CO_2 to cellulose occurs in the natural process of photosynthesis.
It involves the photoelectrochemical reduction of water to hydrogen and oxygen
(Bockris and Tunulli, 1980) and the subsequent formation of CH_2O by chemical steps
from the released H_2 and CO_2. This natural process gives rise to expectations of the
possibilities of "fixing" atmospheric CO_2 by artificial photoelectrochemical means.
The best would be if one could form methanol photoelectrochemically from the
inexhaustible supply of CO_2 in the air.

The first step in such work was the photoelectrochemical conversion of CO_2 to
oxalic acid at an InP electrode using a two-photon cell having $LaCrO_3$–TiO_2 as a
photoanode (Guruswamy, 1979). Decoration of a CdTe electrode with 18 crown 6
ethers in the presence of certain tetraalkylammonium ions has established a photoelec-
trocatalytic effect of the crown ethers which brought about CO_2 reduction largely to
CO, but with small amounts of CH_3OH also formed (Wass, 1989).

Then, using a soluble magnesium anode, it has been possible to demonstrate a standalone photocell that reduced CO_2 to phenylacetic acid (Table 10.1), with the by-production of electricity. The photocathode was p-GaP, which worked according to a Helmholtz model with a high density of surface states (Nakabayashi and Uosaki, 1993). One can see in this kind of work the future possibility of the formation of valuable organics—pharmaceutical compounds—directly from CO_2, with the exciting possibility of photoelectrogenerative fuel cells that would not only produce commercially valuable organics but also feed electricity into a national grid.

10.8.1. Photoelectrochemical Waste Removal

Several photoelectrochemical studies of the photo-oxidation of H_2S have been made. It is possible to use CdS as a photoanode connected to platinum, which evolves hydrogen. Pure sulfur results (Kainthla, 1986). There is an initial chemical reaction:

$$H_2S + 2OH^- \rightarrow S^{2-} + 2H_2O$$
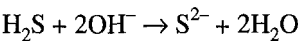
The S^- is then oxidized photoelectrochemically by the receipt of holes at the CdS/solution interface.

$$S^{2-} + 2h^+ \rightarrow S^0$$

$$S^{2-} + S^0 \rightarrow S_2^{2-}$$
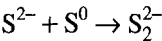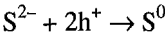
The second chemical reaction leads to the formation of polysulfide and finally sulfur which is a marketable product. Irradiation of CdS with solar light thus continu-

TABLE 10.1
Current Yield of Phenyl Acetic Acid in Various *p*-Semiconductor Synthetic Photocells

Electrode	Current (μA)	Time (hr)	Amount of Product (nmol)	Current Efficiency (%)
p-GaP (50 mM)	3	20	530	44
(2 M)	2	24	350	47
p-GaAs (50 mM)	4	20	630	42
(2 M)	7	16	1000	42
p-InP (50 mM)	0.2	20	—	—

Source: Reprinted from I. Nakabayashi, F. Ushizaki, and K. Uosaki, "A Photoelectrochemical Carbon Dioxide Fixation: Spontaneous Up-Quality Conversion of Organic Compounds," in *Environmental Aspects of Electrochemistry and Photoelectrochemistry*, M. Tomkiewicz, H. Yoneyama, R. Haynes, and Y. Hori, eds. proceedings vol. 93–18, p. 49, Table 1, 1993. Reproduced by permission of the Electrochemical Society, Inc.)

ously produces sulfur and hydrogen. H_2S is a substantial pollutant world-wide that is available from gas wells. Its use as an economic source of hydrogen has good potential.

Correspondingly, many toxic organic compounds (particularly halo-aromatics) have been shown to be oxidizable to innocuous compounds at TiO_2 photoanodes with a counter-electrode of platinum (Bard, 1980; Fox, 1986). In 1980 Guruswamy suggested a design for a colloidal solar reactor for wastes that worked on photoelectrochemical principles.

The greatest waste problem of most developed countries[15] is sewage. The present (chemical and bacterial) means for treating sewage leads to a leftover product, "sewage sludge," and this has to be burned. Incomplete combustion leads to the wide distribution of undesirable products. Alternatively, the sludge can be spread out on land, whereupon rain leaches out toxic products that eventually may reach the water table and drinking supply. Electrochemical sewage disposal is discussed in Chapter 15. Photoelectrochemical processes for such a purpose have a research potential.

10.9. RETROSPECT AND PROSPECT FOR PHOTOELECTROCHEMISTRY, PARTICULARLY IN RESPECT TO THE SPLITTING OF WATER

Like fuel cells (Chapter 13), photoelectrochemistry has had a roller-coaster history. Interest in it began in the nineteenth century (Becquerel, 1839). There were many studies before 1950 (Copeland and Gerrett, 1942) and then the accidental discovery by Fujishama and Honda and its coincidence with what was perceived at the time as "the energy crisis" of the 1970s[16] gave rise to a huge outburst of photoelectrochemical work in the 1980s. Thus, in 1981 and 1982, photoelectrochemically active authors (Bard, Grätzel) were the most quoted in all of physical science.

However, with progress in the fundamental sciences being tied so tightly to the degree of government funding, the sparse Regan years led to a downturn in progress. The 1995 manifesto by Lewis, Tributsch, and Uosaki shows the international nature of the opinion that photoelectrochemistry is a core subject for environmental development in the twenty-first century. Thus, whatever the arguments about the wisdom

---

[15]In the United States and Russia, the most serious waste problem is radioactive and toxic material that has accumulated from weapons production during the Cold War years, closely followed by spent fuel from reactors.

[16]After the Arab–Israeli war of 1973, the Arabs insisted on ownership of all oil under their ground. They raised the price of crude oil by 400%. Yamani, the Saudi-Arabian minister for energy at the time, was quoted as asking western governments which they would prefer, to pay more or have him turn off the Saudi Arabian oil supply; he was cartooned sitting, smiling, his hand on a wheel attached to a gigantic oil pipe. The availability of gasoline was briefly rationed in some western countries.

For about a decade (particularly in the Carter presidency) it was public policy to seek alternative energy sources, and the supply of research funds for photoelectrochemistry was significant. However, President Reagan appointed an ex-dentist as head of the U.S. Department of Energy. More widespread *drilling* for oil was recommended. Funds for renewable research work were cut by about 90%.

of using fossil fuels in the future, it is a fact that their combustion is accompanied by abnormal growth in CO_2. A resulting (and eventually unacceptable) increase in world temperatures seems inevitable unless the growth of atmospheric CO_2 (continuous now since ~1910) is halted. This conclusion by the majority of environmental scientists, not only in the United States, but in all technologically advanced countries, has led to a consensus that solar energy and energy from other renewable (mostly sporadic) resources will have to be made available on a massive scale within the next one to two generations. It is this kind of background that led leading scientists in the photoelectrochemical field to their announcement. Insofar as a solar source is used for energy, photoelectrochemistry becomes a frontier topic.

Along with this background, sociology and politics have made a number of advances in the scientific position that underlies the developments needed in the twenty-first century. These developments have been described above and are only briefly repeated here.

1. The overreliance on the Schottky barrier model for reactions involving adsorbed intermediates must be revised to take into account the high surface state concentration to which they often give rise. This position is emphasized in that the most obvious environmental use of photoelectrochemistry is in splitting water to produce clean hydrogen.
2. Metal islets covering *part* of the electrode surface can sometimes be used to obtain significant increases in rates.
3. The use of cells in which both electrodes are semiconductors and are irradiated improves the efficiency of conversion for most overall reactions.
4. Photoanodes that would otherwise decompose on use can be protected by photoinactive transparent passive layers.
5. Faceting of surfaces to reduce the amount of light lost by reflection should be used in parallel with decorating the surface with catalysts, to increase efficiency.
6. The decoration of the semiconductor surface with dyes should be investigated more fully to ascertain if this approach (which potentially increases the range of wavelength absorbed) can be used to increase the efficiency of cheap TiO_2 in the presence of surface O atoms in the formation of molecular oxygen.

Further Reading

Seminal

1. E. Becquerel, *C.R. Acad. Sci.* **9**: 561 (1939). First paper recording effect of light on electrodes.
2. A. W. Copeland, B. Black, and A. B. Garret, *Chem. Rev.* **31**: 177 (1942). Review of work up to 1942.

3. P. J. Hilson and E. K. Rideal, *Proc. Roy. Soc. London* **199:** 295 (1949). Light effects in metals.

4. J. O'M. Bockris, S. U. M. Khan, and K. Uosaki, *J. Res. Inst. Catal. Hokkaido* **24**:1 (1976). Theory of photoelectrochemistry at metals.

5. W. W. Gärtner, *Phys. Rev.* **116**: 84 (1959). Photoemission from solids; basis of much of Gerischer's subsequent Schottky barrier theory. Origin of neglect of surface properties until 1980s.

6. H. Gerischer and F. Beck, *Z. Elektrochem.* **63**: 500 (1959). Early Schottky barrier paper.

7. M. Green, "Semiconductor Electrochemistry," in *Modern Aspects of Electrochemistry*, J. O'M. Bockris and B. E. Conway, eds., Vol. 2, Ch. 2, Plenum, New York (1959). First formulation of semiconductor electrode kinetics in terms of equations; band bending; surface states; limiting currents.

8. G. C. Barker, *Electrochim. Acta* **13**:1221 (1968). Light effect on metals in the presence of certain agents in solution.

9. H. Gerischer, *Electroanal. Interfacial Chem.* **50**:263 (1975). Photoelectrochemical kinetics.

10. D. S. Ginley and M. A. Butler, *J. Appl. Phys.* **48**:2019 (1977). Photocurrents in the limiting current region interpreted in terms of Schottky barrier.

11. M. A. Butler, *J. Appl. Phys.* **44**: 1914 (1977). Theory of photocurrents in terms of energy gap and flatband potential; transport in rate control.

12. A. M. Bard and B. Krautler, *J. Am. Chem. Soc.* **100**: 4317 (1978). Photoelectrochemical decomposition of wastes.

13. A. D. Beardsley, C. Bookbinder, R. N. Dominey, N. S. Lewis, and M. Wrighton, *J. Am. Chem. Soc.* **102**: 36 (1980). Hydrogen evolution from semiconductors. "Pinned" Fermi levels.

14. J. O'M. Bockris, K. Uosaki, and H. Kita, *J. Appl. Phys.* **52**: 808 (1981). Experimental evidence of surface effects in photoelectrochemical kinetics.

15. R. Memming, in *Comprehensive Treatise of Electrochemistry,* B. E. Conway et al., eds., Vol. 7, p. 534, Plenum, New York (1983). General survey of photoelectrochemistry; Schottky oriented.

16. M. Szklarczyk, J. O'M. Bockris, V. Brusic, and G. Sparrow, *Int. J. Hydrogen Energy* **9**:707 (1984). Substrate effects in photoelectrochemistry.

17. N. S. Lewis, C. M. Gronet, G. W. Cogan, J. E. Gibbons, and G. M. Moddel, *J. Electrochem. Soc.* **131**: 2873 (1984). Nonaqueous solution study of redox reactions at light activated semiconductors confirming applicability of Schottky-type theory.

Modern

1. A. Fujishama and K. Honda, *Nature* **238**: 37 (1972). Claim of first photoelectrochemistry water splitting.

2. J. O'M. Bockris and K. Uosaki, *J. Electrochem. Soc.* **125**: 223 (1977). First theoretical treatment of photoelectrochemical kinetics at high surface source case.

3. E. Buhks and F. Williams, *Proc. Electrochem. Soc.* **82**: 1 (1981). Model for photoelectron transfer from semiconductors.

4. K. Uosaki and H. Kita, *J. Electrochem. Soc.* **128**: 2153 (1981). Tafel lines in photoelectrochemical kinetics when measured at current densities sufficiently below that of the limiting current.

5. Y. Y. Pleskov and Y. Y. Gurevich, *Semiconductor Electrochemistry,* Consultant's Bureau, New York (1986). Textbook.

6. H. Tributsch, in *Modern Aspects of Electrochemistry,* J. O'M. Bockris, B. E. Conway, and R. H. White eds., Vol. 17, Ch. 4, Plenum, New York (1986). Materials for photoelectrodes.

7. M. A. Fox, *Topics Current Chem.* **142**: 75 (1987). Organic photoreactions, some photoelectrochemical.

8. S. Fan and A. J. Bard, *J. Appl. Phys.* **27**:1331 (1988). First use of tunneling microscopy approach to semiconductor surfaces in solution.

9. F. Willig, in *Modern Aspects of Electrochemistry,* R. H. White, B. E. Conway, and J. O'M. Bockris, eds., Vol. 19, Ch. 4, Plenum, New York (1989). Charge transfer at organic crystal surfaces.

10. H. M. Kuhne and J. Schefeld, *J. Electrochem. Soc.* **137**: 548 (1990). A Helmholtz interpretation of the effect of metal clusters; Tafel lines.

11. K. Uosaki, *Trends Anal. Chem.* **9**: 98 (1990). Photoluminescence at electrodes.

12. L. T. Canham, *Appl. Phys. Lett.* **57**: 1046 (1990). First paper on photoluminescence from porous Si.

13. P. V. Kamat, *J. Am. Chem. Soc.* **113**: 9705 (1991). Effect of coatings to protect electrodes.

14. A. Hamnett and R. A. Bachelor, in *Modern Aspects of Electrochemistry,* R. White, B. E. Conway, and J. O'M. Bockris, eds., Vol. 22, Ch. 3, Plenum, New York (1992). Surface states.

15. A. Hagefeldt and M. Grätzel, *Chem. Rev.* **95**: 835 (1995). High efficiency of light conversion using dye-sensitized TiO_2.

16. M. Koinume and K. Uosaki, *Electrochim. Acta* **40**: 1345 (1995). Atomic force microscope studies of GaAs.

17. E. M. Arce, J. G. Ibanez, and T. Mear, *Electrochim. Acta* **40**: 263 (1995). Characterization of the surfaces of photoelectrodes.

18. W. Jaegermann, in *Modern Aspects of Electrochemistry,* R. H. White, B. E. Conway, and J. O'M. Bockris, eds., Vol. 30, Ch. 1, Plenum, New York (1996). Semiconductor/solution interface.

19. K. Uosaki, "Electrochemistry 1992–1995," in *Ann. Report of the Chemical Society,* Vol. 92, pp. 50–58 (1996).

20. M. J. Schimmel and H. Wendt, *Proc. Electrochem. Soc.* **97–20**, p. 16. Anodic formation of tenary semiconductors.

21. D. J. Fermin, E. A. Ponorarev, and L. M. Peter, *Proc. Electrochem. Soc.* **97–20**, p. 62. Electrode processes during the photoillumination of TiO_2.

22. P. Bonkote, P. Compte, and M. Grätzel, *Proc. Electrochem. Soc.* **97–20**, p. 106. Performance of nanocrystalline solar cells.

23. A. Michallis, M. Schweinsburg, and J. W. Schultze, *Proc. Electrochem. Soc.* **97–20**, p. 209. Photocurrent specks on thin films.
24. A. M. Cheperro, M. Alonso-Vonte, P. Salvador, and H. Tributsch,*Proc. Electrochem. Soc.* **97–20**, p. 218. Electroreflectance with Bu_2S.
25. H. Noguchi, T. Kondo, and K. Uosaki, *Proc. Electrochem. Soc.* **97–20**, p. 260. Electroreflectance.
26. S. Licht, P. A. Ramakrishnan, O. Khaselev, T. Soga, and M. Umeno,*Proc. Electrochem. Soc.* **97–20**, p. 358. Multiple band gap cells.

APPENDIX 1. A BRIEF NOTE ON ELECTROLUMINESCENCE AND ELECTROREFLECTANCE

The main topics in photoelectrochemical studies involving semiconductors relate to the absorption of light in certain potential regions, giving rise to the electric currents across semiconductor/solution interfaces. Correspondingly, it is possible to find a potential range at which light is emitted (electroluminescence).

Light emission from porous Si was discovered by Canham (1990). The phenomena are complex and depend on the Si nanocrystal size (Uosaki, 1997).

When monochromatic light is incident upon an electrode, it will be absorbed if the energy of the photons is greater than that of the band gap. The resulting electric current produced depends upon the electrode potential. Correspondingly, a fraction of the light is reflected and the intensity of this light is potential dependent.

Electroreflectance has been studied at the interface BuS_2 (Tributsch, 1997). It passes through a maximum at about 0.3 V SCE.

APPENDIX 2. ELECTROCHEMICAL PREPARATION OF SEMICONDUCTOR ELECTRODES

The electrochemical formation of metallic alloys is well known. As photoelectrochemistry progressed during the 1990s, it became clear that advantages may accrue to semiconducting electrodes containing several different atomic species, and here the possibilities of co-deposition of two or more atomic species are, at least in principle, clear. The accomplishment of such a goal has not been easy because of interfering, unwanted electrode processes, the potential of deposition of which overlaps that of the desired ones.

Three examples of successful processes can be mentioned. The simplest (achieved in the early 1980s by O. Murphy), is CdTe. $Cd\,Se_x\,Te_{1-x}$ (Meisner, 1998), and $CuInS_2$ (Wendt, 1998) are complex semiconductors, the preparations of which have been reported.

Etched Ni substrates are used for the CdTe. The two electrode reactions are:

$$HTe\,O_2 + 3H^+ + 4e \rightarrow Te_{(6)} + 2H_2O \quad E^\circ = 0.31 \text{ (SCE)}$$

$$Cd^{2+} + Te_{(s)} + 2e \rightarrow Cd\,Te_{(s)} \qquad E^\circ = 0.13 \text{ (SCE)}$$
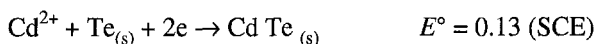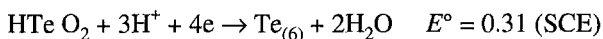
The potential has to be controlled very precisely. If it is insufficiently negative, the Cd/Te ratio of 1:1 is not obtained. If it is too negative, dendritic deposits of Cd appear. The more complex compounds (see above) are grown anodically (e.g., Cu_2S co-deposited with $In_2 S_3$).

To what extent the films grown electrochemically have decisive advantages over those grown with other techniques is not clear yet. However, one can see that great variety (e.g., ternary and quaternary alloy formations) is possible, and the availability of potentiostatic control and nonaqueous solutions may be helpful.

APPENDIX 3. HIGH-RESOLUTION TECHNIQUES IN THE STUDY OF SEMICONDUCTOR SURFACES

Micrometer and atomic (nm) resolution can be obtained in the study of semiconductors in their function as electrodes. The two most prominent techniques used at the end of the century were scanning laser spot (SCS) and scanning tunneling microscopy.
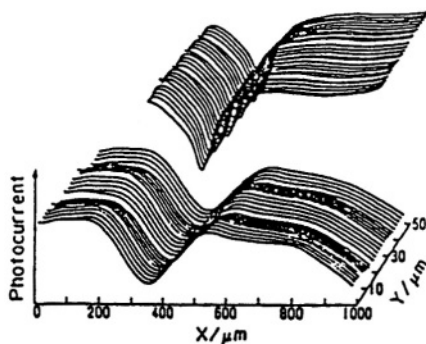
Fig. 10.30. A line plot showing the laser spot photocurrent across a step on an InSe electrode before (lower lines) and after Pt treatment (upper lines). (Reprinted from K. Uosaki, D. Carlson, H. Kita, and B. Holmstrom, in *Electrochemistry in Transition*, B.E. Conway, O.J. Murphy, and S. Srinivasan, eds.) Plenum, 1992, p. 453.)
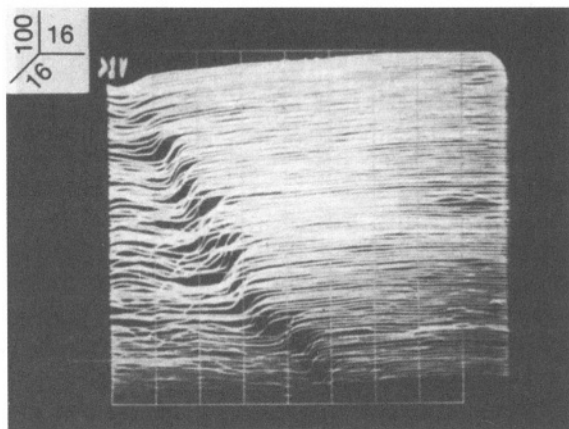
Fig. 10.31. Image of p-Si in air; V_{bias} = +1.0 V (Reprinted from M. Szklarczyk, A .Gonzalez-Martin, O. Velev, and J. O'M. Bockris, "STM Studies of p-Si(111) substrate in Air and in Electrolytic Environment," *Surf. Sci.* **237**, Copyright 1990, Fig. 7, with permission from Elsevier Science.)
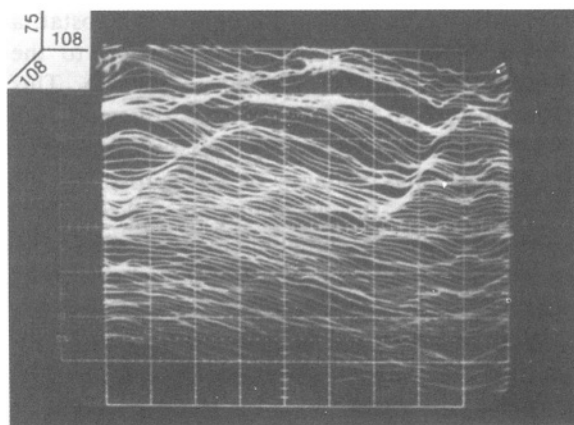
Fig. 10.32. STM image of p-Si in solution; V_{bias} = −1.0 V. Electrode potentials: E_{Si} = −1.0 V, E_{tip} = 0.0 V (Reprinted from M. Szklarczyk, A.Gonzalez-Martin, O. Velev, and J. O'M. Bockris, "STM Studies of p-Si(111) Substrate in Air and in Electrolytic Environment," *Surf. Sci.* **237**, copyright 1990, p. 305, with permission from Elsevier Science.)

The presence of scratches, grain boundaries, and steps reduces the photocurrent because these defects encourage hole–electron recombination.

The scanning laser spot technique allows the lateral resolution of the photoelectrochemical properties of the semiconductor and permits the visualization of the defects (Furtak and Parkinson, 1980). Figure 10.30 shows the results of sweeping and laser spot over a step in InSe. A step is a recombination center and the recombination of holes and electrons increases.

The resolution obtained with the laser spot technique is far exceeded by scanning tunneling microscopy where, in some cases, atomic resolution in electrochemical cases has been reached (Szklarczyk and Velev, 1989). The first successful studies of semiconductors in air (Fig. 10.31) and in an electrochemical situation (Fig. 10.32) were made on p-Si 111 (Gonzalez-Martin, 1990). It was found that the electrochemical formation of SiO_x and SiOH induces surface states at 0.25 V above the valence band at the surface (Fig. 10.33).
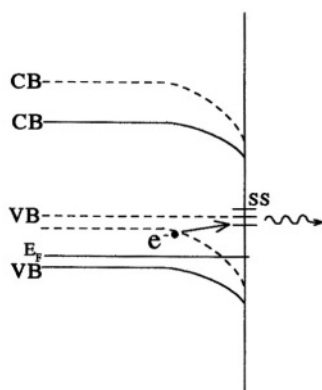
Fig. 10.33. Energy band diagram for negative biasing of semiconductor sample in solution. Solid lines, band before biasing; dashed lines, bands after biasing; SS, surface states. (Reprinted from M. Szklarczyk, A. Gonzalez-Martin, O. Velev, and J. O'M Bockris, "STM Studies of p-Si(111) Substrate in Air and in Electrolytic Environment," *Surf. Sci.* **237**, copyright 1990, Fig. 12, with permission from Elsevier Science.)

The importance of the surface state configuration for the operation of semiconductors is stressed by these studies (Uosaki and Koinuma, 1992).

EXERCISES

1. Calculate the equivalent of 1.5 eV photons in (a) the wavelength of light and (b) its frequency and its wave number. (Bockris).

2. Consider the solar spectrum. (a) Identify the fraction of solar light with photons of energy greater than 3 eV, 1.6 eV, and 1.0 eV. (b) Which of these photons is toward the UV and which toward the IR? (Bockris)

3. With the sun overhead in a cloudless sky, the power of sunlight falling on 1 square mile of land is slightly over 1 kW. In fact, the maximum solar energy falling on the earth varies with latitude and is generally substantially less than this. In addition, for fixed solar collection (i.e., collectors that do not track the sun), the amount of collectible solar light varies with the time of day (and of course the weather). The solar energy received in a given location has to be measured over a number of years and the average taken. The best collecting areas are clearly near the equator and in desert areas.

   Calculate the area of land needed for solar energy for a community of 10 million people (city dwellers, using 10.6 kW of power for all purposes—domestic, transport, industrial, and military) using the following basic assumptions: efficiency of conversion of solar to electrical energy; 20%; average amount of solar energy falling on collecting area, $1/2$ kW m^{-2} for 10 hr per day. (Allow 15% extra area for service pathways to allow maintenance work on the collectors.) (Bockris)

4. Draw the potential–distance relations for a semiconductor without surface states: (a) for an n-type electrode acting as a photoanode and (b) for a p-type electrode acting as a photocathode. (Bockris)

5. Although the "Schottky barrier" model (negligible surface states) is applicable for some electrochemical reactions involving redox species and electrode reactions with no surface bonding of intermediate radicals, most practical, useful photoelectrochemical reactions involve significant numbers of surface states. Draw the potential–distance relations for the corresponding Helmholtz approximation: (a) for a photocathode and (b) for a photoanode. (Bockris)

6. When solar light falls on an electrode semiconductor (n-type) such as TiO_2 ($E_g \sim 2.9$ eV), the only photons to be absorbed (and hence to give rise to holes in the valence band) are those with energy in excess of E_g. From a diagram

of the solar spectrum, calculate the fraction of solar energy that TiO_2 will adsorb. (Bockris)

7. The maximum energy conversion efficiencies are as follows:

TABLE E.1

Electrode	p-InP	p-InP-Co	p-InP-Ni	p-InP-Au	p-InP-Pt
ε_{max} (%)	1.8	5.4	14.9	19.2	22.8

Interpret this. (Contractor)

8. A surface of p-InP was modified by the electrodeposition of submonolayer amounts of various metals and the photocurrent vs. potential behavior studied. The photocurrents observed at 0 V vs. NHE for various surface treatments are shown in the Table E.2. Calculate the maximum efficiency of energy conversion in each case and comment on the observed trend. The light source was an Xe lamp and incident light intensity was 50 mW cm^{-2}. (Contractor)

TABLE E.2
Dependence of Photocurrent on Surface Treatment with Various Metals. Currents were measured in 1 M H_2SO_4 with a platinum counter-electrode

I_{ph} (mA cm^{-2})	0.73	2.19	6.04	7.81	9.27
Electrode	p-InP	p-InP-Co	p-InP-Ni	p-InP-Au	p-InP-Pt

9. The following data were obtained during interfacial capacitance measurements of a single-crystal n-TiO_2 electrode in 0.1 M TBAP (tributylammonium phosphate) + CH_3CN at a frequency of 500 Hz. Calculate the flatband potential on n-TiO_2 in this electrolyte and the concentration of majority carriers. Assume $\varepsilon = 86$.

TABLE E.3

E (V vs SCE)	0	−0.1	−0.2	−0.3	−0.4	−0.5	−0.6	−0.7	−0.8	−0.9	−1.0
C (µF)		0.285	0.30	0.32	0.34	0.37	0.40	0.44	0.49	0.56	0.60	0.61

Calculate the carrier density. (Contractor)

10. By studying the temperature dependence of the carrier concentration in a semiconductor sample, several things can be learned about its properties. Figure E10.1 shows the free carrier density ($n + p$) of a semiconductor sample measured as a function of temperature and plotted in a "useful" form. Based on the data in this plot, you should be able to discern:
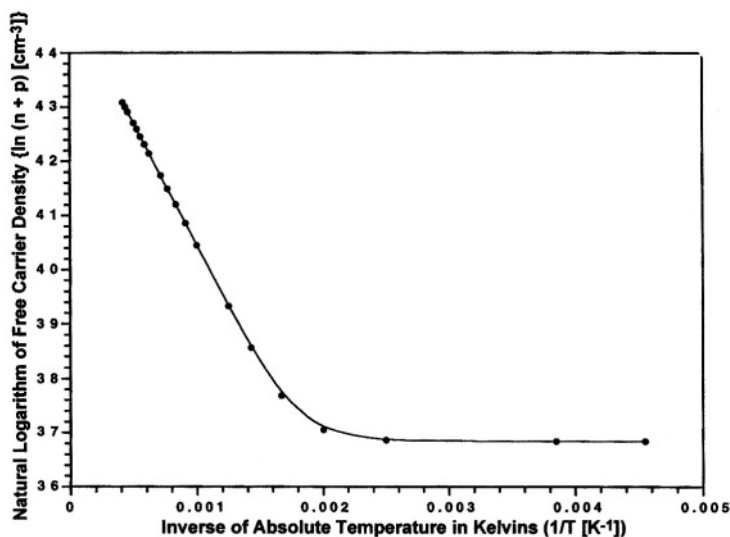
Fig. E10.1.

(a) the dopant density of the sample, assuming that the dopants are all of one type and that they ionize completely in the temperature range under consideration and (b) the band gap of the semiconductor. (Lewis)

(Hint: This is *not* a "real" semiconductor; you will not find anything that matches its E_g!)

11.   Calculate or otherwise determine the following parameters for the five semiconducting materials Ge, Si, GaAs, CdS, and TiO_2. (The required data are Table E.4, which may be incorporated into the problem or an appendix.)

(a) Calculate the intrinsic carrier concentration (n_i) at room temperature (300 K). (b) Give an example of an atom substitution that would provide each type of dopant (donor, acceptor) in these semiconductors. (Note that in compound semiconductors the atom substituted for is just as relevant as the identity of the atom that replaces it, i.e., in GaP, Si in a P site acts as an acceptor, but Si in a Ga site acts as a donor!) (Lewis)

12.   (a) What density of extrinsic dopants would be needed to double the total free carrier (electron + hole) concentration from its intrinsic value in a sample of silicon at 300 K, assuming complete ionization of the dopants? (b) In a sample of TiO_2 at 300 K? (Lewis)

13.   (a) How thick would a wafer of Si have to be to absorb 99% of the 800-nm photons hitting it, assuming no reflection losses at the front surface and complete

TABLE E.4
Semiconductor Data
(KEEP THIS TABLE! YOU WILL NEED IT FOR ON SEVERAL HOMEWORK SETS!
The data are precise for undoped semiconductors at 300 K.)

Physical Property	Semiconductor				
	Ge^a	Si^a	$GaAs^a$	CdS^b	$TiO_2^{\ c}$
Band gap (E_g) (eV)[d]	0.70	1.12	1.42	2.42	3.2
Gap transition type[d]	Indirect	Indirect	Direct	Direct	Indirect
Electron mobility (μ_e)(cm^2/V · s)	3600	1350	−6750	340[a]	0.4
Hole mobility (μ_p)(cm^2/V · s)	1800	480	300		16
Electron effective mass, longitudinal ($m_{e,1}^*/m_e$)	1.64	0.98	0.067	0.21[a]	~10^3
Electron effective mass, transverse ($m_{e,t}^*/m_e$)	0.082	0.19	—	—	—
Hole effective mass, light ($m_{p,1}^*/m_e$)	0.044	0.16	0.082	0.80[a]	~0.8[e]
Static dielectric constant (κ or ε_s) (dimensionless)	16.0	11.9	13.1	5.4	114[e]
Electron affinity (χ) (eV)	4.13	4.01	4.07	4.5[f]	
Conduction band effective state density (N_c) (cm^{-3})	1.0×10^{19}	2.8×10^{19}	4.7×10^{17}	2.4×10^{18}[g]	7.9×10^{20}[g]
Valence band effective state density (N_v) (cm^{-3})	6.0×10^{18}	1.0×10^{19}	7.0×10^{18}	1.8×10^{19}	1.8×10^{19}
Mass density (ρ_m) (g/cm^3)	5.3267	2.328	5.32	4.82	4.23
Atomic density (ρ_a) (g/cm^3)	4.42×10^{22}	5.0×10^{22}	2.21×10^{22}	2.0×10^{22}	2.19×10^{22}
Crystal structure[d]	Diamond	Diamond	Znblende	Wurtzite	Anatase
Breakdown field (ε_{crit}) (V/cm)	~10^5	~3×10^5	~4×10^5		~6×10^4
Intrinsic Debye length (L_D) (μm)	0.68	24	2250		

Source: R. Rossi, Stanford University, 1998.
[a]S. M. Sze, *The Physics of Semiconductor Devices,* 2nd ed., Wiley; New York, 1981.
[b]CdS exists in two distinct crystal forms: hexagonal and cubic. The data given here refer to the hexagonal phase, the more stable of the two crystal forms.
[c]TiO$_2$ exists in two distinct crystal forms: rutile and anatase. The data given here refer to the anatase phase in colloidal form, since it is most commonly used in semiconductor applications.
[d]M. X. Tan, P. E. Laibinis, S. T. Nguyen, J. M. Kesselman, et al., *Prog. Inorg. Chem.* **41**: 21–144 (1994).
[e]B. Enright and D. Fitzmaurice, *J. Phys. Chem.* **100**:1027–1035 1996.
[f]J. W. Mayer and S. S. Lau, *Electronic Materials Science: For Integrated Circuits in Si and GaAs*, Macmillan, New York, 1990, p. 16.
[g]These values were estimated assuming only one equivalent energy minimum in the respective bands; they are at best reliable to an order of magnitude. (Rossi)

reflection at the back surface? (b) What thickness of GaAs would be required to accomplish the same feat? $(\alpha_{Si})_{800nm} = 1 \times 10^3$; $(\alpha_{GaAs})_{800nm} = 9 \times 10^3$. (Lewis)

PROBLEMS

1. The maximum possible energy that a photoelectrode can absorb is all radiation striking it, the photons of which have energy $< E_g$ if *one* photoelectrode is used in a photocell and the counter-electrode is a metal. This is, then, the energy available for conversion to new materials by a photoelectrochemical process.

   Using tables of the E_g's of semiconductors, illustrate the advantage of using photocells in which two electrodes are irradiated. In your quantitative discussion, explain why such an arrangement may lead to a higher efficiency for the conversion of solar light to materials than the conventional use of a single electrode. (Bockris)

2. In photoelectrochemical work, it is usual to plot only the highest values of the photocurrent as a function of potential. This is necessarily an S-shaped curve, with the highest values of the photocurrent eventually being controlled by diffusion of carriers inside the semiconductor. (A thermally activated electrode reaction on a metal has a similar shape near the limiting current caused by transport of ions to the electrode.)

   Consider the photoelectrochemical kinetics as a function of potential at photocurrent densities that are 1% of the limiting photocurrent and less. What will be the nature of the photocurrent electrode potential in these regions? Explain your reasoning. (Bockris)

3. In the Schottky barrier approximation for photoelectrodes, virtually all the potential differences near the semiconductor solution interface lie inside the semiconductor. However, for photoelectrodes that evolve H_2 and O_2 (i.e., are photo water splitters), the H and O adsorbed on the surface of the semiconductor cause surface states for electrons there. In such a "high surface state" (Helmholtz) approximation case, the potential difference around the semiconductor/solution surface moves out into the solution and the potential difference in the semiconductor is greatly reduced in extreme cases, becoming negligible.

   Calculate the potential difference in the Helmholtz layer in the solution for a semiconductor in which the surface state density is 10^{14} cm^{-2} (assume the effective dielectric constant is 6 and the thickness of the double layer 5 Å). As an approximation, neglect the contribution due to oriented dipoles of adsorbed water. (Bockris)

4. Photoelectrocatalysis is the term given to the positive effect on the rate of photoelectrochemical reactions that occurs in some cases when the semiconductor surface is decorated with metal islets covering a small fraction of the surface. Consider the following facts for the photoevolution of hydrogen from water.

The shift in potential (ΔE) at a given current density on the i_{ph}–V plot caused by the decoration of the surface by metal islets is linear with the $\log i_0$ of the corresponding thermal reaction on the metals concerned. ΔE increases with an increase in the work function of the decorating metal. The heat of activation for the photoevolution of hydrogen from water on a metal/decorated surface is the same as that for the thermal evolution of H_2 on the same metal.

Interpret these facts to favor the view that photoelectrocatalysis occurs because of a rate-controlling reaction at the metal/solution interface, or, alternatively, that it occurs because the metal added to the surface shifts the Fermi level of the semiconductor. (Bockris)

5. The photosynthetic reaction in nature that creates biomass represents an attractive model for photoreactions, i.e., one would like to be able to take CO_2 from the atmosphere and "fix it"—convert it to a useful product using the energy of solar light.

Using the available data, discuss whether an approach to this societally valuable project would be best approached: (a) directly, i.e., by attempting to find the photoelectrocatalysts that would lead from CO_2 directly to MeOH (the most desirable product); or (b) by using photoelectrocatalysis to split water and then attempting to use the adsorbed H (present as an intermediate in the reduction of water) to react chemically with CO_2 to form methanol. Draw appropriate reaction schemes to illustrate your ideas. (Bockris)

6. Economic photosplitting of water is clearly a worthwhile goal, for it could lead to a supply of clean fuel. (a) Discuss as quantitatively as you can whether direct photosplitting could provide economic advantages (one plant) that could compensate for the cost reduction achieved by the higher efficiency of the photovoltaic production of electricity, followed by normal water electrolysis.

In contriving optimal photosplitting of water, it is desirable to irradiate *two* photoelectrodes. In considering photoanodes, it is difficult to achieve suitable properties with semiconducting oxides, which tend to be stable under oxygen evolution. On the other hand, arsenides and sulfides appear to have more suitable photoelectrochemical properties, but are clearly unstable when exposed to O_2 evolution. (b) What experimental procedures could be used to overcome this difficulty? (Bockris)

7. A variety of experimental measurements can be used to determine free carrier densities (and thus dopant levels) in semiconductors. Four-point probe resistivity measurements are the most common because they are relatively painless, nondestructive, and can be performed on thin wafers. There is more to them than meets the eye, however.

It might seem that given a semiconductor sample of adequate size and shape, a "direct" measurement of its resistivity could be made, by multiplying the resonance made by an ohmmeter by the length of a silicon sample. (a) Give three
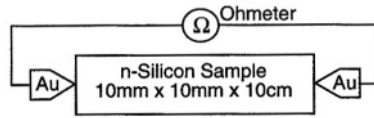
Fig. P10.1.

reasons why this is not possible using the setup in Fig. P10.1.

Typically, one works with thin, round semiconductor wafers, ranging in size from 5 inches to less than a quarter inch in diameter, and measures their resistivities using a four-point probe setup. Suppose you use a four-point probe to measure the resistivity of an n-type (p-doped) Si sample. A micrometer tells you that the average thickness (t) of the 30-inch diameter wafer is 0.408 mm. The spacing on the probe is 0.050 inch. You collect the following current-voltage data at room temperature:

TABLE P.1

Current Through Outer Probes (µA)	Potential Difference Across Inner Probes (µV)
10.62	349
21.90	590
29.77	757
41.37	1002
49.00	1162
60.71	1409
70.39	1613
79.35	1799
90.47	2032
99.91	2228
110.11	2440
120.78	2658
130.73	2861
140.50	3060

(b) Determine the free carrier density.

8. The Fermi–Dirac and Maxwell–Boltzmann statistical distribution functions are widely used in semiconductor physics, with the latter commonly used as an approximation to the former. The point of this problem is to make you familiar with these distribution functions: their forms, their temperature dependencies, and under what conditions they become interchangeable. Throughout this problem, use the energy of silicon's valence band (E_{vb}) as the zero of your energy scale.

(a) Consider a perfectly intrinsic sample of Si. Calculate the location of the Fermi level (E_F) in this material at 0.001, 150.0, 300.0, and 600.0 kelvins. Be

reasonably precise; include the effective density of state correction, but assume that Maxwell–Boltzmann statistics apply.

(b) Plot the expected electron occupancy (the probability of finding an electron in a state at energy E) in the silicon at each of the temperatures above. Remember that electrons obey Fermi–Dirac statistics. (i) Plot the expected occupancy on a linear scale from –0.05 to +1.05 for energies between –0.2 and +1.4 eV; use a linear scale for the energy (x) axis. (ii) Plot the expected occupancy on a logarithmic scale ranging from 10^{-10} to 1.0 for energies between –0.2 and +1.4 eV; again, use a linear energy (x) axis.

(c) Plot the Maxwell–Boltzmann distribution function at the same four temperatures for the same silicon sample. (i) Plot the Boltzmann distribution function on a linear scale; use the same axis limits as you used in part (b) so that you can compare the two plots directly. (ii) Plot the Boltzmann distribution function on a logarithmic scale, again using the axis limits used in part (b) so that you can compare the two plots directly.

(d) Compare the two distribution functions. (i) At what temperatures and over what energy ranges are the Fermi–Dirac and Boltzmann distribution functions appreciably different? (ii) Are the two distribution functions distinguishable under any of the conditions considered here for energies in the conduction band, where states actually exist? Under what conditions might the two distributions differ appreciably for energies in the conduction band? (Lewis)

9. When gallium arsenide (GaAs) is grown by molecular-beam epitaxy (MBE), Si dopants are specifically incorporated into Ga sites; when it is grown by organometallic vapor-phase epitaxy (OMVPE), Si dopants are specifically incorporated into As sites. Suppose that you have two samples of GaAs, both doped to 1.3×10^{16} cm^{-3} with Si, but one grown by MBE and the other grown by OMVPE. Where will the Fermi level (E_F) be in each case, relative to the position of the nearest band (valence or conduction) in GaAs? (Lewis)

MICRO RESEARCH PROBLEMS

1. Consider a clean sample of n-GaAs (free of surface states) brought into contact with a redox/active solution so that a junction with a stable reproducible barrier height of 0.90 V is obtained. Plot the behavior of the following quantities as a function of the dopant density over a dopant concentration range of 10^{15} to 10^{17} cm^{-3}: (a) the width (W) of the depletion region in the semiconductor, in nanometers; (b) the value of the maximum electric field strength (E_{max}) in the GaAs; (c) the number of electrons (Q_{eq}) transferred across the interface in forming the junction, and the direction in which they moved.

(Note: It is safe to make several approximations and assumptions in working out this problem and in solving semiconductor physics problems in general. First,

go ahead and use the depletion approximation. Second, assume that all the dopants ionize completely. Third, assume that the potential drop occurs completely across the less polarizable of the two phases involved in forming a contact; polarizability increases in going from insulators to semiconductors to electrolytic solutions to metals. Finally, assume that the surface state density is negligibly small unless you are warned otherwise.) (Lewis)

2. In polycrystalline semiconductor samples, the excited-state lifetime of electron–hole pairs is so short that photocurrent collection is efficient only for carriers created within the space charge (depletion) region. Thin-film processes offer an inexpensive way to prepare large solar arrays, but the semiconductors formed by such processes are almost inevitably polycrystalline. It is not wise to use semiconductor films thicker than the depletion width in such devices because the additional thickness contributes only extra grain barrier boundaries for the majority of carriers to surmount on their way to the back contact. The additional thickness does not provide any additional photocurrent.

    (a) For the case of polycrystalline n-GaAs doped to 6.8×10^{16} cm^{-3} in contact with Au (Φ_b for this junction is 0.75 V), calculate how thick a layer of GaAs would be required to collect 99% of the light, and compare this with the width of the depletion region. (Because the sample is polycrystalline, assume that no reflection occurs at the back contact; consider the incident light to be monochromatic at 800 nm.) (b) What do your results in part (a) mean with regard to the maximum external quantum yield we might expect from such a thin-film GaAs device? (Lewis)

This page intentionally left blank

CHAPTER 11

SELECTED ASPECTS OF ORGANOELECTROCHEMISTRY

11.1. INTRODUCTION

11.1.1. The Modernization of an Ancient Subject

The beginning of bioelectrochemistry (Galvani, 1791) is sometimes said to mark the beginning of the study of electrochemistry. However, it was organoelectrochemistry that started only shortly after Galvani's adventures with the frog's legs. Von Arnim is quoted as the originator. Volta (1800) had shown that an electric current driven by his battery of silver and zinc plates flowed through a solution. Von Arnim was romantically inclined, a poet. How fine, he thought, were he to take the good French wine he drank and give it an extra "kick" by passing this novel electric current through it. The result was not the one he intended, for the wine now tasted vinegarish. He was told that he must have produced acetic acid (1801).[1]

The nineteenth century was a time of intense rivalry between Germany and England for industrial supremacy (i.e., new technology). When the great Michael Faraday, at the Royal Institution in London, considered Von Arnim's attempted challenge, he decided that a suitable reply to this German impertinence would be to use Volta's electric current to *decompose* Von Arnim's acetic acid.

This was in 1836, and for another century, organoelectrochemistry can truly be said to have stumbled about. At this early time, there was no Butler–Volmer equation to show that for a given organoelectrochemical reaction (and no other) to continue, one had to keep the *electrode potential* the same. Were the potential to change, the nature of the organoelectrochemical reaction might change and hence the products would be a mixture of more than one compound: a mess. Two big changes in this field

---

[1]This piece of electrochemical lore comes from Dr. Norman Weinberg.

occurred during the twentieth century, and these have made organoelectrochemistry a lively field, just as it was at the beginning. The first of these was one of the seminal advances in electrochemistry, an electronic feedback device (the potentiostat) introduced by Hickling in 1942. Whenever the potential of the anode grew more positive than the intended value, the device signaled an outside power source to reduce the anodic current through the working electrode, bringing the potential down again to the desired value. The device acts similarly in respect to any cathodic wandering of the potential, so that the potential (and the product of the now constantly maintained electrode reaction) remains constant.

However, a second change was necessary before there could be a successful electrochemical industry. Carrying out reactions at an electrode is, after all, a two-dimensional affair and one measure of the economics of a process is the amount of the product produced per unit volume. Could an electrode be made that was three dimensional? It was Fleischmann and Goodridge (1970s) who (independently) produced somewhat different calculations showing the feasibility of a 3D electrode of certain (limited)[2] dimensions. The modern organoelectrochemical industry was on its way.

11.1.2. The Plus and Minus of Using an Electrochemical Route for Synthesis

It is much easier to produce the correct energy level for the receipt of electrons from a reactant into the electrode in an anodic synthesis than to arrange for effective electron transfer in a homogeneous solution. In the electrochemical method, one calmly adjusts the potentiostat setting in a solution that may be near room temperature. In solution, one has to worry about adjusting reactants and products—probably catalysts—until it "works." And it may be that the chemical reaction won't work at a significant rate until the temperature is raised.

Again, the products of an electrochemical oxidation are the product itself and hydrogen; this is not pollutive, for the hydrogen can be burned in air and the resulting water returned to the cell.[3] Correspondingly, reduction will consist of the product desired and evolved oxygen. The corresponding chemical process is not under tight

---

[2]A simple model of a 3D electrode is a sponge made of metal and full of solution. However, the immediate thought that this would mean a very large internal area available for reaction is far from the truth. The reactants tend to get held up in diffusing in from the solution bulk. Much of the inner part of the sponge cannot react because of solute exhaustion. Worse than this, unless the solution is *very* concentrated in electrolytes, the lengthy passage through the pores creates an IR drop, and this loss of interfacial potential difference for points inside the sponge greatly reduces the local current density so that frustratingly, 3D electrodes are only as it were 21/2D; i.e., they do increase the effective area from which the current can be drawn but this happens much more near the mouth of the pores, not far inside (Section 13.6.3).

[3] A better way to do it is to use an oxygen cathode and *reduce* oxygen to water. This greatly decreases the cell potential and hence the electricity costs. The water used up at the cathode is replaced. If one *burns* H_2 to do this, one loses the heat energy produced.

control as to reaction pathway offered by a potentiostat, so the reaction "wanders" and may produce unintended side reactions, some of which are likely to reach the atmosphere.

Apart from the ease of precise control in an electrochemical path to synthesis, there is the unique feature of being able to force the electrode reaction to take place against its own ΔG. This is because the principal rule of *chemical* equilibria is $\Delta G = 0$, but in electrochemical equilibria, the equilibrium condition is $\Delta G = -nFE_{rev}$. Thus, if the cell potential is exactly E_{rev}, the chemical reaction in the cell is at equilibrium and nothing happens. However (in contrast to what can be done chemically), moving the potential of the working electrode in a more negative direction than its reversible potential stimulates the reaction to take off in a cathodic direction at a fixed rate; i.e., it acts to reduce the reactant:

$$R + ne \rightarrow P$$
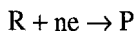
Correspondingly, if the potential of the working electrode is moved in the positive direction compared with its potential at equilibrium, the reaction begins to occur anodically, i.e., in oxidative mode:

$$P \rightarrow R + e$$
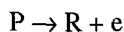
Thus, in an electrosynthesis, one has the ability to make the reaction do what one wants and at a chosen speed (though one might get hemmed in by the presence of potentials of undesired reactions, and of course by the limiting currents of the processes concerned.

Electrocatalysis of organoelectrochemical reactions is a subject on its way; it needs another decade or two of research funding to get enough knowledge so that the field can be developed rationally. However, quite a few catalysts for organoelectrochemical reactions are already here (some are shown in Table 11.1).

As to the downside of electrosynthesis, when a company considers a new synthesis, one may have to admit that organic synthesis is a huge edifice and electroorganic reactions, as known so far, are a small part of the whole. One of the reasons for this is that the ratio of organic chemists to electro-organic chemists in the United States is more than 10:1.

There is one alleged difficulty that is only apparent. It refers to the fact that energy in the form of electricity is 2–3 times more expensive than energy in the form of heat. However, the cost of the electrical energy needed per unit weight of product is generally less than 10% of the whole cost of the product. Furthermore, one cannot make a chemical process go backward against its free energy except by electrochemical means, so that here the comparison is the cost of doing something compared with the cost of not doing it.

TABLE 11.1
Electrode Materials for Electro-Organic Reactions

Type	Conventional	Nonconventional
Metal	Pb, Pt, Hg	Cd-steel, Cd-Cu, Ni, Pt-Ti, Pt-C
Metal oxide	PbO_2-Pb, Fe_3O_4(cast)	PbO_2-Ti, Fe_3O_4, NiOOH-Ni, TiO_2-Ti (cathode only), Cr_2O_3-Ti (anode only), TiO-Ti, RuO_2-Ti, MnO_2-Ti, VO_2-Ti (cathode only)
Carbon	C (electrographite, Acheson graphite)	Carbon black-filled polymers, graphite-filled plastics, graphite-felt, glassy carbon (anode only), porous carbon, poly-p-phenylene (synthetic metal)
Metal compound	$FeSi_x$, SiC	TiC-Ti, TiN-Ti, TiS_2-C, BN

Source: Reprinted from F. Beck, in *Ulmann's Encyclopedia of Industrial Chemistry*, Vol. 49, p. 235, 1985, with permission from Wiley-VCH.)

Finally, when it comes to electricity cost, electro-organic reactions do not *have* to be driven backward; many desired products can be obtained by electrochemical cell reactions that have negative ΔG's. If these reactions can be broken up into two reactions in an electrochemical cell, they form a kind of fuel cell that is intrinsically electrogenerative, so now electricity is being made and can be used elsewhere—the electricity costs have turned into a gain (Section 13.3).

These remarks must be balanced by some characteristic difficulties of using the electrochemical path. Sometimes, and in spite of tight potential control, two or more reactions take place at the same time and give not one product but a mixture. Correspondingly, overoxidation may occur; the intended oxidation may continue to a further step by means of a chemical driving force outside the control of the potentiostat.

11.2. DETERMINING THE MECHANISMS OF ORGANOELECTROCHEMICAL REACTIONS

11.2.1. Introduction

Determination of the successive steps and discernment of the rate-determining step in an organochemical reaction is a more complex task than settling the mechanism of the simpler and basic reactions historically quoted, e.g., hydrogen evolution or oxygen reduction. The first question that has to be answered before one begins on a detailed study of a reaction mechanism is: Are the phenomena being observed the results of a single overall reaction or are they mixed up with the products of reactions that are close in potential?

Another question concerns the relevant time domain for the investigation. It is always simpler to carry out the reaction as quickly as possible, i.e., to carry out potential sweeps or even cyclic voltammetry (Section 8.6). Among the reasons for this is that the catalyst surface may undergo a degree of deactivation in minutes, let alone hours or months. Polymerized and largely unreactive side products ("gunk") may build up on the electrode surface at the longer times (weeks, months) in which an industrial electrochemical reaction must work without attention.

Thus, in potential sweep work, when the electrode potential is changed too quickly, the various intermediate radicals of the reaction will not correspond to the θ's of the steady-state functioning of the reaction and the information obtained, and hence the mechanism determined from it may be irrelevant in providing the information needed for, say, the design of electrocatalysts to last for 3–6 months of continuous production with the reaction in the steady state.[4] Thus, there is not much point in carrying out an "academic" mechanism investigation (mostly done with fast sweeps) because the time domains may be too short for radical buildup to the steady state.

Correspondingly, organic-electrochemical reactions sometimes offer so many alternative paths (at least as seen by simple stoichiometric considerations) that the best that can be done is to reduce the dozen or more possible rate-determining steps that formalistic consideration seems to show are possible, down to two or three alternatives. The degree of detail with which one may determine a mechanism is always a matter of what is available, the instrumentation (both electrochemical and spectroscopic), time, and funds.

Analysis of mechanism is constantly improving. The simultaneous use of several spectroscopies (ellipsometric, FTIR, Raman) in parallel may increase the power to detect intermediate radicals and their surface kinetics. Combining such means with pattern recognition techniques (Section) may eventually lead to fast determination of organic reaction pathways and rate-determining steps. Around the beginning of the new century, however, it is more a matter of combining a good knowledge of organic chemistry with signals obtained by using a half-dozen different electrochemical approaches, aided by what spectroscopy can reveal about the entities present on the surface of the electrode in the steady state.

The example discussed in the next section is one of intermediate complexity (anodic oxidation of β-cyanoethyl ethers). A simpler example would not give the flavor of a "real" investigation.

11.2.2. Anodic Oxidation of β-Cyanoethyl Ethers

Wermeckes and Beck examined the anodic oxidation of β-cyanoethyl ethers, $ROCH_2CH_2CN$ (R = Me, Et, t-Bu, CH_2CH_2CN), at Pt and PbO_2 electrodes in 1.6 M

---

[4]Organic substances adsorb quite slowly (~10^3 s to equilibrium) on noble metal surfaces. This is by no means due to the organic molecules making their way through the forestlike platinized platinum surface. For example (Uosaki, 1996), 2-mercaptohydroquinone takes about 15 min to reach constant concentration and potential to adsorb to equilibrium on an evaporated gold layer.

H_2SO_4. They found that most of the product was cyanoacetic acid (CEA), with the other product being the hemiacetayl (Hac) acid corresponding to the R group (i.e., for R = Me, formic; for R = Et, acetic). When R = Et and the starting material was β-$EtOCH_2CH_2CN$ (EPN), Hac was found at 81% current efficiency (low temperature), but at high temperatures CEA was found at a current efficiency of only 31%.

Experiments on the mechanism of the oxidation of β-cyanoethyl ethers were carried out by Wermeckes and Beck in both a one-compartment and a two-compartment cell. In the electrolysis of ethylpropionitrile (EPN) there is an early appearance of β-oxypropionitrile, which is an intermediate in the formation of $NCCH_2CH_2OCH_2CH_2CN$ and acetaldehyde. Evidently Et is being attacked at the CH_2 group to give a hemiacetal in a two-electron oxidation (Fig. 11.1). This is split in solution to yield the intermediates mentioned. HCN is found as a side product, and it may come from an anodic attack on the CH_2 next to the cyano group. Cyanohydrin would be formed, and this is readily saponified. Wermeckes and Beck made sure that the cleavage of the C–O bond occurs only under anodic electrochemical conditions, that is, that it is not a solution reaction.
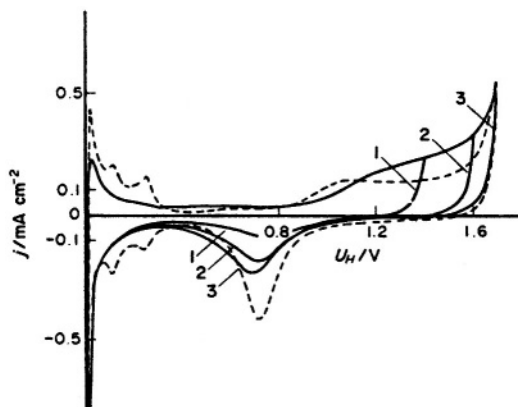
Fig. 11.1. Cyclic Voltammetry on Pt electrodes in 1.0 $M\,H_2SO_4$. Voltage scan rate, $v_s = 100\,mV\,s^{-1}$. Variation of potential of positive scan reversal: ---, basic curve; –, curve in the presence of 1 M EPN. Voltages of scan reversal U_H: (1) 1.4 V, (2) 1.6 V, (3) 1.7 V. Shown is the sixth cycle after starting the experimental sequence. Freshly prepared Pt electrode for the first sequence was at the most positive potential of scan reversal. T = 20 °C, N_2 purging. (Reprinted from B. Wermeckes and F. Beck, *Electrochim. Acta* **30**:1491, copyright 1985, with permission from Elsevier Science.)

What is the site of an anodic attack, and is this a function of the distribution of the electron density in the molecule? In $Et-O-CH_2-CH_2-CN$ (note the overall reaction equation),

$$\overset{3}{CH_3}CH_2O\overset{2}{CH_2}\overset{1}{CH_2}CN + 3H_2O \rightarrow CH_3COOH + NCCH_2COOH + 8H^+ + 8e \quad (11.1)$$
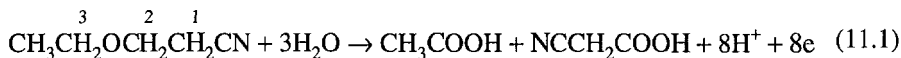
the CH_2 group 3 has a higher δ^- owing to hyperconjugation effects at the methyl group. The methylene group 2 has the second highest electron density. The reactions leading to the hemiacetal would be:

$$R-CH_2-O-R' + H_2O \longrightarrow R-\underset{\underset{OH}{|}}{CH}-O-R' + 2H^+ + 2e \quad (11.2)$$

$$R-\underset{\underset{OH}{|}}{CH}-O-R' \overset{[H^+]}{\longrightarrow} RCHO + R'OH \quad (11.3)$$
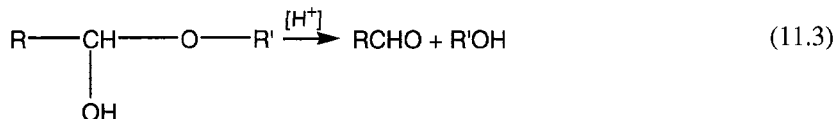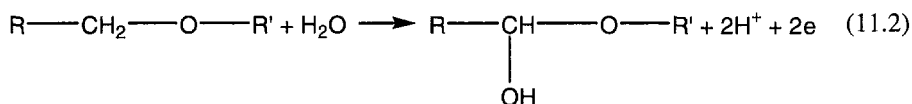
The Tafel lines show unusually steep slopes ($b = 240-300$ mV dec^{-1}), and Wermeckes and Beck interpreted them as due to strong adsorption of depolarizer molecules on Pt (see Fig. 11.2). It is interesting to note that the cyclic voltammograms (Fig. 11.1) *show that organic molecule adsorption leads to a reduction of O coverage on Pt.* Correspondingly, impedance measurements show the decrease in capacity expected if there is significant organic adsorption. The model that these workers suggest is reproduced in Fig. 11.2; in this model, the organic and the oxide are present together, the organic occupying the edge sites at which O_2 evolution usually occurs. Thus, oxidation to the hemiacetal proceeds according to Eq. (11.2).

If the reaction proceeds via an electron transfer from the plane *RP* (see Fig. 11.2), one could see a factor (*f*) in the Tafel equation, so that:

$$i = Ke^{fFV/2RT} \quad (11.4)$$

and if $f \approx 1/2$, $b \approx 240$ mV. The overall mechanism is summarized in Fig. 11.3.

11.2.3. The Manufacture of Nylon

Nylon is a synthetic material that has largely replaced silk in clothing. The electrochemical part of the synthesis originated in Russian work (Knunyants, 1954), but the process was developed in the United States by Baizer (1964). Knowledge of the molecular mechanism is due partly to teams led by Bard (1974) and partly to those led by Savéant (1975).

The first step in the synthesis is the polymerization of acrylonitrile to adiponitriles

Fig. 11.2. Schematic representation of the interface at a platinum anode for anodic oxidation of EPN. (Reprinted from B. Wermeckes and F. Beck, *Electrochim. Acta* **30**:1491, copyright 1985, with permission from Elsevier Science.)

$$H_2C = CH_2CN \xrightarrow[2e]{2H^+} NC - (CH_2)_4 - CN$$

acrylonitrile   adiponitrile

Figure 11.4 is helpful in understanding the mechanism.

There are two scenes shown in the figure. In one, acrylonitrile (AN) is adsorbed on the electrode in the presence of (on the left), Li^+; on the right, tetraethylammoniun, NEt_4^+. The latter is the chosen scene. On the one hand, its potential for discharge of NEt_4^+ is even more negative than that of Li^+ and this allows the potential of the electrode to be taken to more negative values than would otherwise be possible (because

Fig. 11.3. Wermeckes and Beck's reaction mechanism for the anodic oxidation of the β-cyanoethyl ether EPN to the main products Hac and CEA. (Reprinted from B. Wermeckes and F. Beck, *Electrochim. Acta* **30**:1491, copyright 1985, with permission from Elsevier Science.)
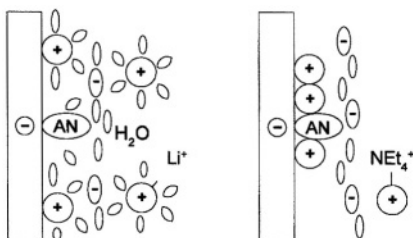
Fig. 11.4. Reduction of water activity at the electrode/electrolyte interface in the electrodimerization of acrylonitrile (AN). Water is repelled from the interface by NEt_4^+ and this decreases the likelihood of hydrogen evolution competing with the acrylonitrile reduction. (Reprinted from F. Beck, *Erdöl Kohle Erdges*, [Oil, Coal, Natural Gas] *Petrochem.* **41**:287, 1988.)

competing discharge of the supporting electrolyte would take electric current away from production of the final product. Another advantage is having NEt_4^+ in the double layer. This ion acts hydrophobically (Vol. 1, Section 2.20.6) and is shown in Fig. 11.4 as expelling water from the double layer (because it adsorbs strongly). Such an expulsion of water is an advantage because it diminishes the probability of OH_2 discharge to form H_2—an undesirable side reaction, again increasing the probability of electrons being used for the reduction of acrylonitrile.

The electrochemical mechanism arising from the work by Bard and, independently, by Savéant, leads to an eec mechanism:

$$X–CH{=}CH–X + e \rightarrow X–CH–\overset{\cdot}{C}H–X^-$$

$$2X–\overset{\cdot}{C}H–CH–X^- \xrightarrow{\text{Net}_4^+} \overset{(-)\quad NEt_4^+}{(X–CH–CH–X)_2}$$

The way the latter entity is written bears the following message: This is an anionic polymerization and can be understood as possible because of the efficient screening of the anion repulsion by the NEt_4^+ cation.

The Tafel slope is $2RT/F$ and the first of the above reactions is therefore rate determining. A unity value for the reaction order both with respect to acrylonitrile and adiponitrile helps lead to the following mechanism (Beck 1986):

11.3. CHIRAL ELECTRODES

11.3.1. Optical Activity at Electrodes

It will be seen (Section 11.6.1) that metal electrodes can be usefully modified by adsorbing on them organic compounds of a structure suited to some specific task. For example, if a Pt electrode has adsorbed upon it a monolayer consisting of 4,4′-dipyridyl, then

proteins can undergo rapid electron exchange reactions on the modified electrode. In the absence of the modifying layer, large proteins are apt to fall to pieces upon adsorption.

A particular class of modified electrodes consists of those containing a layer of asymmetric compounds, and such electrodes are termed "chiral." If one uses these electrodes in organic synthesis, the compound produced may also be asymmetric and optically active. One of the better-known examples of such phenomena is called the *Sharpless process* (Finn and Sharpless, 1986; Katsuki, 1996). In such processes, the electrode is modified by asymmetric compounds that lead to epoxidation and dihydroxylation of olefinic compounds, but in an asymmetric form. An example is shown in Fig. 11.5, in which the hydroxylation occurs either on the top or the bottom of the enantiomorphic surface.

The first work of this kind was carried out as long ago as 1936 (Criegee). It was shown that osmium tetraoxide could lead to the *cis*-dihydroxylation of olefins. The electrochemical method of making such compounds is simple and more cost effective than competing chemical methods.
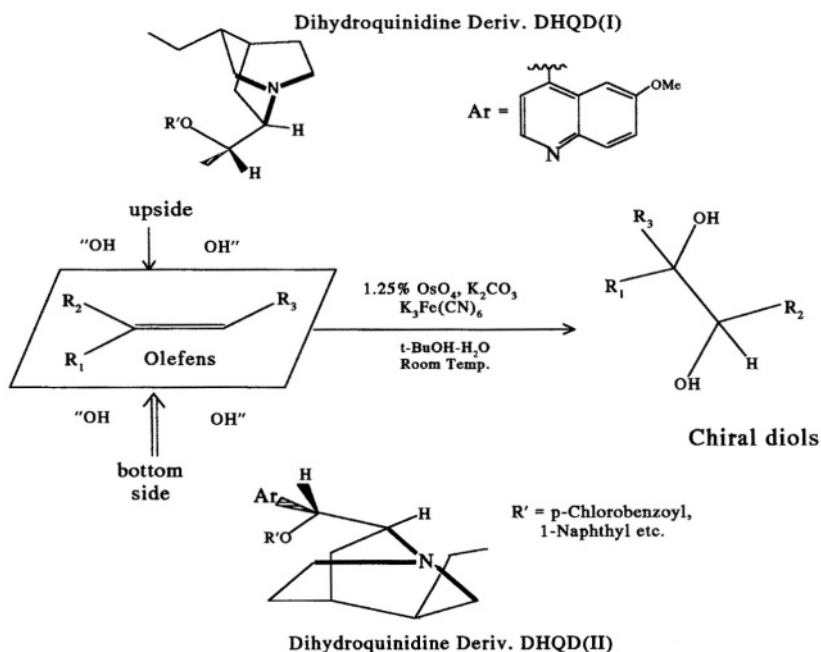
Fig. 11.5. Os-catalyzed asymmetric dihydroxylation of olefins by enantiomeric differentiation (chiral induction). (Reprinted from S. Torii, "Electrochemical Asymmetric Syntheses of Chiral Diols and Epoxides" *Interface*, Winter 1997, p. 46, Scheme 1. Reproduced by permission of the Electrochemical Society.)

11.4. ELECTRO-ORGANIC SYNTHESES

11.4.1. Cell Design

Electrochemistry in organic syntheses began in the later 1800s. The best-known early synthesis is that of Kolbe in which semiesters of dicarboxylic acids dimerized the alkyl chain to form long-chain hydrocarbons. As already mentioned, progress in the industrial development of electro-organic syntheses was by no means rapid until the introduction of electronic means for maintaining an electrode potential (hence the reaction) constant (Hickling, 1942) and that of the packed bed (a 3D electrode) (Goodridge, 1969).

In 2000, about 110 chemicals were being produced by electro-organic syntheses at a rate of more than 10,000 tons/year. The best-known method has already been presented in this chapter (Section 11.2.3): it is the electro-hydrodimerization of acrylonitrile to adiponitrile as part of the synthesis of nylon.

Several improvements in cell design have occurred since 1980. The fraction:

$$\frac{\text{Moles (hr)}^{-1} \text{ produced}}{\text{Cubic meters of plant kilowatts}}$$

is a measure of the goodness factor of an organic synthesis. It is desirable to reduce the cathode–anode distance toward zero to minimize the IR losses and thereby to reduce the kilowatts needed. This is particularly necessary in nonaqueous systems in which the electrolyte resistance tends to be far higher than in aqueous solutions. A few modern cell designs are listed in Table 11.2.

For cells that lack IR loss-increasing membranes, the Swiss-roll cell design (Ibl, 1975) may be the answer. Two metal foils act as electrodes and are separated by two insulating plastic masks that act as turbulence promoters. This cell gives large area and low IR. It is shown in Fig. 11.6.

TABLE 11.2
Modern Concepts of Cell Design

Cell	Ohmic Voltage Drop	Mass Transfer Rate	Specific Electrode Area
Cells with turbulence promoters	High	High	Small
Rotating cylinder electrode cell	High	High	Small
Trickle-tower cell	High	High	Large
Pump cell	Low	High	Small
Swiss-roll cell	Low	High	Large

Source: Reprinted from G. Kreysa, in *Ulmann's Encyclopedia of Industrial Chemistry*, Vol. 49, p. 200, 1985, with permission from Wiley-VCH.)
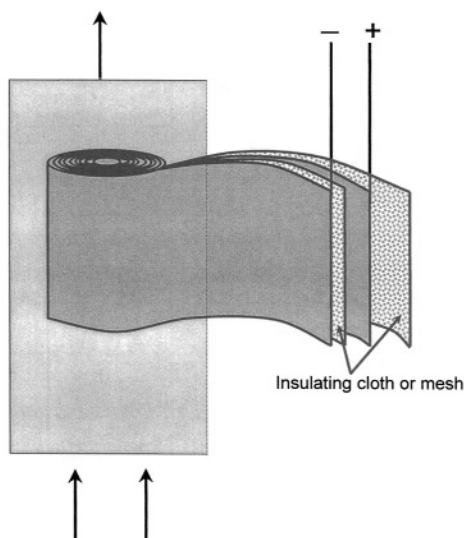
Fig. 11.6. Swiss-roll cell. (Reprinted from G. Kreysa, in *Ulmann's Encyclopedia of Industrial Chemistry*, Vol. 49, p. 200, 1985, with permission from Wiley-VCH.)

11.4.2. New Electrode Materials

A number of new electrode materials have also come into use in organoelectrochemistry. One aim is to avoid expensive noble metals and electrodes that dissolve or decrepitate on use. Thus the use of carbon electrodes in chlorine production in the electrolysis of brine has been supplanted by the so-called DSA (dimensionally stable anodes) electrodes. These are essentially TiO_2 electrodes, but this material is modified by the addition of other oxides (e.g., RuO_2) to make it conducting and to increase its electrocatalytic properties. The advantage of the use, of such electrodes is indicated by their name; they do not change shape in use, whereas the carbon electrodes in earlier chlorine cells had a restricted lifetime because of their gradual disintegration and change in shape upon use.[5]

---

[5]The history of the development of the DSA electrodes is a classic of its kind. The original inventor, Henri Beer, was of the traditional type; his work started in a small private laboratory. The first offer of the rights to his invention was made by a large British chemical concern for $25,000. Successful patenting was concerned with the help of deNora of Milan, Italy. World patents have been taken out by Diamond Shamrock of the United States.

Dimensionally stable anodes represent, from the commercial point of view, one of the more important inventions in the history of industrial electrochemistry. Even so, carbon electrodes were an entrenched and stable part of the mindset of the chlorine production community. deNora's sales technique was at first to offer use of the new electrodes at zero cost, only asking for half the dollar gains obtained by a company's use of the new electrodes. This technique created a market for the new electrodes, so that it could soon be replaced with a more normal sales technique.

Among electrodes coming into practical use in the past two decades are $Ni\text{-}NiO_x$; $Ti\text{-}Cr_2O_3$; Ti_4O_7 and $Ti\text{-}VO_x$. Ebonex is a name referring to a certain brand of reduced TiO_2 (the reduction is necessary to give nonstoichiometry and as a result sufficient electronic conductivity). Doping of TiO_2 by Sb_2O_3 has also sometimes been used to increase conductivity.

11.4.3. A Moving Frontier

Weinberg and Mazur (1997) have reviewed recent advances in organic synthesis. Among the new syntheses named by them (but not yet commercialized) are the following:

- *Selective Organic Electrofluorination.* The electrogeneration of fluorine from the electrolysis of a mixture of KF and HF has been well known for many decades. Organic perfluorination using HF and HF/KF is used widely. Now the use of HF complexes with tetraalkylammonium fluoride has led to the anodic difluorination of dithioacetals.
- *C–C Bond Formation.* Electrochemical generation of radicals (e.g., carbanions) has been shown to give rise to stereoselective intramolecular cyclization.
- *Chiral Electrode Formation.* This was described in Section 11.3.
- Electrohydrodimerization of formaldehyde to ethylene glycol:

$$2CH_2O + 2H^+ + 2e \rightarrow HO\text{--}CH_2\text{--}CH_2\text{--}OH$$
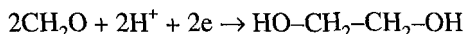
- Dimethyl maleate is being electrochemically converted to 1,2,3,4-butane tetracarboxyllic acid.
- Anthraquinone is being made at pilot plant scale from anthracene. The $Ce^{3+/4+}$ couple is used with methane sulfonic acids. The steps involve anodic oxidation of Ce^{3+} and the use of Ce^{4+} outside the cell to convert naphthalene to napthaquinone, which is then converted to anthraquinone via a step involving butadiene.

It is worth noting that no electrogenerative process[6] leading to organic synthesis on an industrial scale has as yet been publicized. Eventually such processes may need a substantial section to themselves in an electrochemistry text.

11.5. ELECTRONICALLY CONDUCTING ORGANIC POLYMERS

11.5.1. Introduction

The general (classical) concept in chemistry is that metals are highly conducting electronically; solutions of dissolved salts conduct significantly by ionic movement,

---

[6]Namely, a process with a negative ΔG (i.e., a spontaneous reaction), broken down into two electrode reactions and producing electricity plus the product compound (see Section 13.3).

but other compounds have negligible electrical conductance; if some conductance is present, it may be due to the action of impurities.

This view began to be doubted in the later 1950s and by the mid-1960s it had been realized that dozens and perhaps hundreds of inorganic substances (silicon, gallium arsenide) could be *made* to conduct by electronic "doping," i.e. by adding a tiny (e.g., 10^{-8} mol cm^3) concentration of "impurity" atoms, designed to give electrons to the nonconducting form or to accept them from it (hole conduction) (see Section 6 .10).

To Shirakawa (1971) we owe the discovery of *ionic* doping. The compounds concerned are organic polymers and arise from the polymerization and ionic doping of five organic compounds: acetylene, pyrrole, aniline, phenylene, and thiophene (Section 4.9).

Suitably made polymers of these substances[7] have an oxidized and reduced form, and the two forms have dramatically different characteristics. The reduced form is a semiconductor like many others. It is, however, the conductance of the oxidized form that is so impressive, for its conductance may exceed 10^4 Si, i.e., it enters the range of values in which metals conduct.

The usefulness of organic compounds that are highly conducting cannot be emphasized too strongly. Of course, a "doping" step is necessary before these remarkable properties can be observed. In this sense, the electronically conducting polymer (the archetype of such a substance is polypyrrole) are similar to semiconducting inorganic compounds, which must also be doped before the high conductance switches on. However, the mechanism of doping for the organic compounds is radically different from that of substances such as Si and GaP. Instead of dissolving small traces of electron-denying or -accepting impurities into the compounds to increase conductance by many orders of magnitude (as with, e.g., CdSe), the polymers (e.g., polyaniline) are made electronically conducting by doping them with ions. Thus, the undoped polymer is immersed in a solution of, say, $1M$ LiClO$_4$. Here, it is made a cathode and Li$^+$ enters the interstices of the polymer (intercalation) and causes its remarkable conductance (polyacetylene doped with Li$^+$ has a specific conductivity of 10^3 mhos cm^{-1}).

By the late 1980s it seemed that electronically conducting polymers offered a new electrochemistry. The five classes of polymers have been mentioned as giving rise to electronic properties. However, by substitutions of side groups in its structure each polymer may give rise to many new compounds, so that the five electronically conducting groups could well eventually yield dozens of new electrode materials, each with its own surface structure and electrocatalytic properties. The ability to fine tune electrocatalysis seemed to be in sight.

---

[7]The preparation of polymers is largely electrochemical. Great care is needed to ensure very pure starting materials (e.g., aniline, pyrrole). A doping agent [e.g., K$_3$Fe(CN)$_6$] is necessary. One then follows a cyclic voltammogram (Paik, 1992), which indicates the amount of polymer formed—its thickness—in the successive cycles.

Experience has taught caution as to the amount of research needed to develop the possibilities of this new field. The compounds "switch on" as conducting electrodes at a certain potential as one moves in the anodic direction. Each polymer has a *range* of potentials in which it is stable and one must not go too far, for if one advances the potential too far in the anodic direction, the polymer will undergo irreversible oxidation. However, if the electrode process to be run operates within the range of stability for the polymer, the development of new electrodes may become a reality.

Another limiting factor is that it may not be possible to use some of these electronically conducting organic compounds in aqueous solutions. This is true of the polymer most used in new polymer batteries, polyacetylene, for which an appropriate solvent is propylene carbonate (see Fig. 4.115).

Nevertheless, the discovery that high electronic conductance could be induced in organic polymers is a most important event in electro-organic chemistry, and progress in research to develop the electrochemistry of such compounds (which began in the 1980s) is still far from reaching the exponential part of its S-shaped plot.

11.5.2. Ionically Doped Organic Polymers as Semiconductors

The best description of conducting organic polymers from the point of view of solid-state physics, and one very relevant to an understanding of the electrochemistry, has been given by Schulze (1993). He used poly-3-methyl thiophene on a gold base and carried out absorption spectroscopy, XPS, and ultraviolet photoelectron spectroscopy (UPS) on the reduced form to obtain the energy levels of the substance regarded as a semi conductor. He found a photocurrent when the polymer-coated electrode was irradiated at negative potentials (Chapter 10). The doped oxidized form, however, no longer behaves as a semiconductor.

Thus, the *reduced* form of poly-3-methyl thiophene is an intrinsic semiconductor and the Fermi level lies between the valence band and the conductivity band. The effect of oxidation is to introduce a surface state in the band gap between π and π^* orbitals. The Fermi level is decreased when the compound loses electrons, and metallic properties appear when an increasing number of electrons build a new but only half-filled band. These situations are shown in Fig 11.7.

Thus, from a solid-state chemistry point of view, the conducting polymer poly-3-methyl thiophene is in the reduced state, a semiconductor with a band structure. Intercalating with ions and oxidizing makes the compound behave as a metal from 0.45 to 1.1 V on the NHS.

11.5.3. General Properties of Electronically Conducting Organic Polymers

11.5.3.1. *Status of Polypyrrole.* Polypyrrole (Diaz, 1981) has been examined and analyzed more than other electronically conducting organic polymers. It is the archetype of this class of compound.
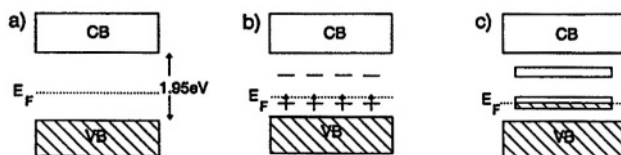
Fig. 11.7. Band model for different degrees of oxidation. (a): Reduced state, (b) partially oxidized, (c) metalliclike state. (Reprinted from K. Stöckert, B. Kessel and J. W. Schultze, "Absorption, Photocurrent and Photoelectron Spectra of Heterocyclic Polymers," *Synth. Met.* **41–43**:1296 copyright 1991, Fig. 1 with permission from Elsevier Science.)

11.5.3.2. *Use of Polypyrrole in Electrocatalysis.*   This material is useful as a bed for, e.g., small-sized Pt particles. The increase in the rate per gram of Pt deposited results from the small size of the Pt particles, which remain stable on polypyrrole, probably because Pt binds to polypyrrole with sufficient strength to prevent diffusion and cluster formation in which many small particles gather together to form one particle (many atoms now being lost to catalytic action because they no longer live on the surface. The polyaniline surface itself has been used in the reduction of CO_2 (Blajeni, 1984) and as a protective film for the use of n-Si as a photoanode (Contractor, 1984).
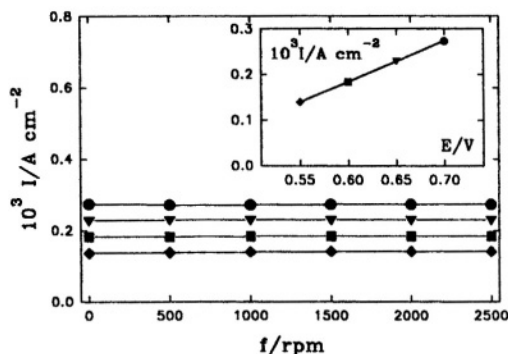
Fig. 11.8. Current density as a function of rotation rate during growth of electrodeposited polypyrrole film at 0.550 (◆), 0.600 (■), 0.650 (▼), and 0.700 (●) V. (Reprinted with permission from D. J. Fermin, M. Mostany, and B. Scharifker, "Electronically Conducting Polymers: Synthesis and Electrochemical Properties of Polypyrrole," *Curr. Topics Electrochem.* **2**:132–136, 1993.)

11.5.3.3. *The Oxidation and Polymerization of the Monomer.* The effects of stirring and studies of absorption spectra (Fig. 11.8) measurements indicate that oligomers—incipient polymers of various member (e.g., nine) chain lengths—gather near the electrode surface. These baby polymers get their start from the activity of radical cations created by monomers at the electrode surface. The growth to adult polymers from the oligomers involves nucleations, which can be interpreted from the maxima in the current–time transient (Fig. 11.9). The rate of the buildup is inhibited by stirring.

The characteristics and structure of the polymers (the length of the polymer and the distance between rings, cross linking, etc.) can be determined spectroscopically. Thus (Watanabe, 1981) a polypyrrol is made up of chains joined through carbon (see Fig. 11.10 for quinguepyrrole).

Electro-oxidation is the key to the introduction of metallic properties (Fig. 11.11). In the conventional theory, polarons[8] with spin and nonspin bipolarons (positive charges) are the result of oxidation. They are attractive to anions, which diffuse in from the solution and up into the interstices of the polymer until the charges on its surface are balanced. This charge generation in the polymer is the slow step in forming a conducting polymer. In polypyrrole, the single polarons exceed the bipolarons.

11.5.4. The Structure of the Polypyrrole/Solution Interface

11.5.4.1. *Relevant Facts.* Microscopic examination of the surface of ionically doped polypyrrol shows a fractal surface. The real surface area can be probed by using organic compounds (e.g., *p*-nitrophenol) of various sizes and finding, by UV-visible measurements of the change in solution concentration caused

---

[8]A polaron is to electrostatic energy what a phonon is to vibrational energy (Section 4.9.2). Originally the term applied to a conduction electron in a salt lattice. It is the electron and its electrostatic interactions with the surroundings together with the strain energy that is called a polaron. More generally, it represents the quantized energy of a charge in a lattice.

There are some (Heinze, 1996) to whom the polaron "explanation" of the ionic introduction of electronic conductivity in organic compounds is specious! The roughness factor of 400 would limit the degree of penetration of ions into the interstices of the polymer. However, Li^+ or even ClO_4^- is of course much smaller than the test molecules (large dye molecules) which are generally used to probe the real area. Thus, one might conceive of a model of the polymer that is "all fibers," the intercalation being all pervasive. It is obvious that an Li^+ ion adsorbed on the surface of a fiber will promote an electron that may indeed be free to move under a field, i.e., to conduct.

Such models do not seem to explain the high specific conductivity observed in electronically conducting compounds. In an alkali metal, there is one conducting electron per atom. If some electronically conducting polymers are to conduct to within 1 or even 10% of this, it would seem to require 0.01 or 0.1 conductivity electrons per atom, and that is difficult to visualize as a consequence of surface adsorption of ions, which will seldom exceed $\theta = 0.1$ for surface occupancy. The mechanism by which such adsorption stimulates conductance *inside* the fibers has not yet appeared in understandable form.

by their adsorption, the molecules adsorbed per geometric area of the surface. The roughness factor for ion-doped polypyrrol in contact with an aqueous solution (e.g., of nitrophenol) was found to be 465 (pore radius 4.8 Å).

Nyquist plots (imaginary vs. real impedance) were matched against a variety of equivalent circuits (Miller, 1992). The best fit was with equivalent circuit 7 in
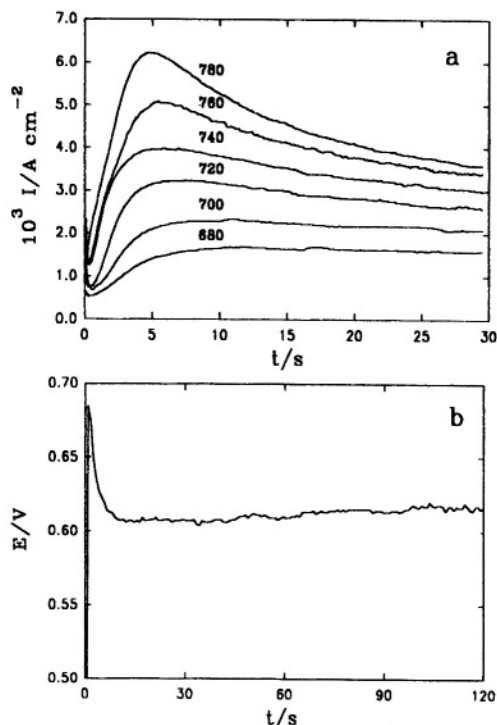
Fig. 11.9. (a) Current transient for the formation of polypyrrole films on gold electrode from a 28.9×10^{-3} M solution of pyrrole in aqueous 1 MKNO$_3$ at the different potentials (mV) indicated. (b) Galvanostatic deposition of polypyrrole on gold electrode from a 50×10^{-3} M solution of polypyrrole in aqueous 1 MKNO$_3$ at a current density of 1 mA/cm^2. (c); Cyclic voltammogram of vitreous carbon in 0.0036 M pyrrole and 1 M KNO$_3$ aqueous solution at 100 mV/s. (Reprinted with permission from D. J. Fermin, M. Mostany, and B. Scharifker, "Electronically Conducting Polymers: Synthesis and Electrochemical Properties of Polypyrrole," *Curr. Topics Electrochem.* **2**:132–136 (1993).)
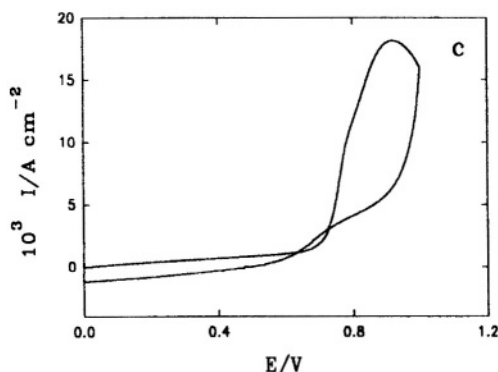
Fig. 11.9. Continued

Fig. 11.12. The results of Hall measurements of mobility are shown in Table 11.3. The Mott–Schottky plot showed a flatband potential of –0.23 V on the NHS. Some electrode kinetic measurements (Miller, 1992) are shown in Fig. 11.13.

The ΔE value between peaks (see Chapter 8) is about 0.1 V, indicating a fast electron exchange between polypyrrol and quinone-hydroquinone in solution. However, a decrease in velocity is shown in the steady-state reaction visible in the Tafel lines of Fig. 11.14.

11.5.3.2. *Structure.* A semiconductor model (Miller 1992) that fits the facts outlined is shown in Fig. 11.15. The material acts as an intrinsic semiconductor on the negative side and a metalized semiconductor on the positive side. The bands shown in the figure are bent upward because the flatband potential was negative with respect to the potential range in which polypyrrol is oxidized and highly conducting. It follows that in the semiconducting state, there is an accumulation of holes at the interface (Chapter 7).
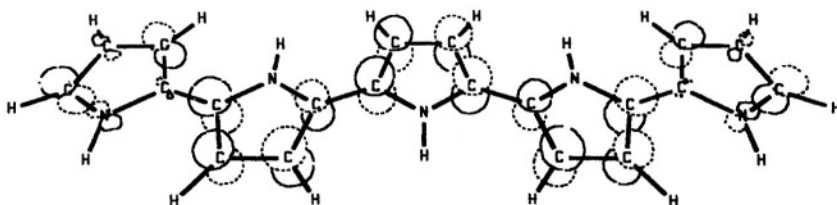
Fig. 11.10. Three-dimensional representation of the singly occupied molecular orbital (SOMO) for the quinquepyrrole. (Reprinted with permission from D. J. Fermin, M. Mostany, and B. Scharifker, "Electronically Conducting Polymers: Synthesis and Electrochemical Properties of Polypyrrole," *Curr. Topics Electrochem.* **2**:132–136, 1993.)
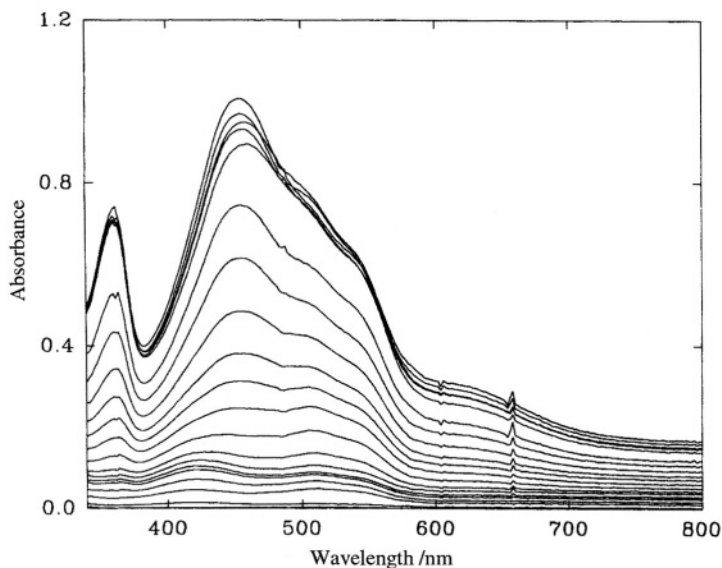
Fig. 11.11. Evolution of the visible spectrum during anodic oxidation of 36 mM pyrrole in aqueous 1 M KNO$_3$ solution on stainless steel grid electrode at 0.7 V (SCE). Spectra were obtained at 60-s intervals. (Reprinted with permission from D. J. Fermin, M. Mostany, and B. Scharifker, "Electronically Conducting Polymers: Synthesis and Electrochemical Properties of Polypyrrole," *Curr. Topics Electrochem.* **2**:132–136, 1993.)

According to the best-fit equivalent circuit (Fig. 11.13, circuit 7), the surface states represent the easy path for electron entry and exit from the double layer region at the semiconductor/solution interface. The double layer is connected linearly to surface states that are in series with the space charge region. Owing to the high carrier concentration and the presence of surface states, the space charge region is only 10 Å in thickness.

As to an interpretation of the drop of catalytic activity with an increase in pH that is observed, it is likely that the nitrogen atoms of the polypyrrol ring would be the active sites for charge transfer (Fig. 11.16). However, with the increase in pH, the proton on the these atoms would be lost, i.e., the activity per electron exchange would be reduced (Fig. 11.17).

11.5.3.3. *Practical Electrochemical Uses of Electronically Conducting Polymers* (see also Section 4.9.2). A large volume of work examining applications of electronically conducting polymers is now available and the details can ben found in the reading list at the end of this chapter. The possibilities of using

Fig. 11.12. Equivalent circuits considered for use in modeling the polypyrrole/electrolyte interphase. (Reprinted from D. L. Miller and J. O'M. Bockris, "Structure of the Polypyrrole/Solution Interphase," *J. Electrochem. Soc.* **139**:970–975, 1992. Reproduced by permission of The Electrochemical Society, Inc.)
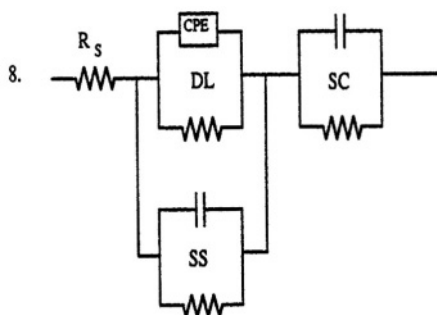
Fig. 11.12. Continued
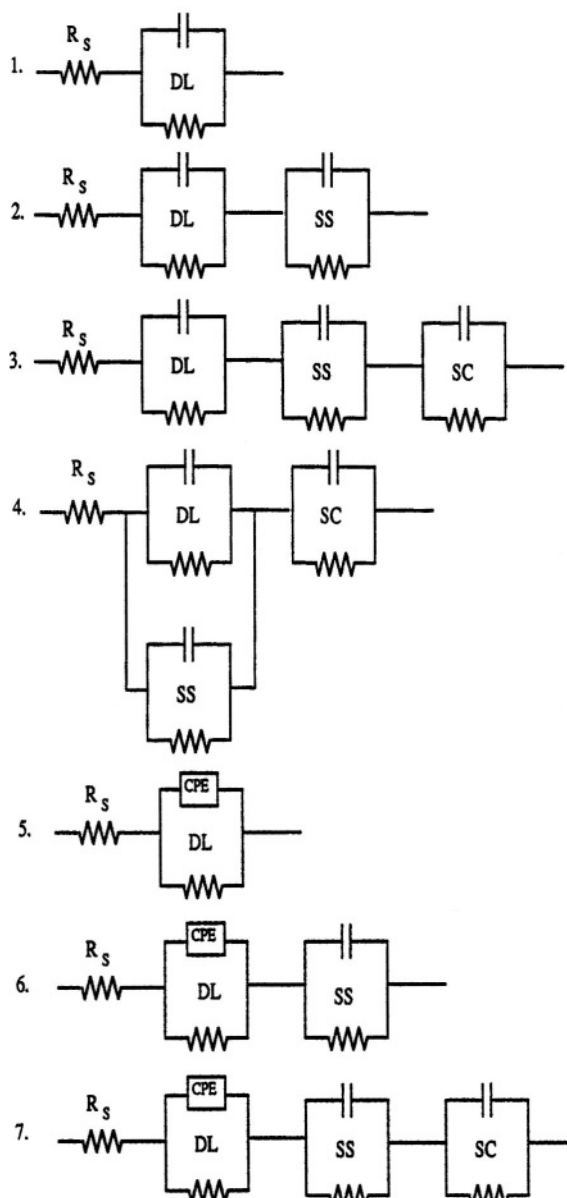
electronically conducting polymers as electrode materials seem large because one could almost endlessly change surface structure in a controlled way. Microelectrode tips could be rationally constructed from a material such as polypyrrole and specific properties introduced (see Fig. 11.18).

The use of these materials as *beds* for microparticles of metals has been successful. A more ambitious idea would be to use the structure of the polymer surface itself and the skills of the organic synthesis chemist to supply groups with specific structures, i.e., to *design* catalysts (see Section 11.6.3).

Activity in the battery area using electronically conducting polymers has been intense (see Fig. 11.19). The possibilities (MacDiarmid, 1997) point to the provision of cheap (because of the relative cheapness of organic compounds) batteries, e.g., for massive use as a power source for bicycles. The counter-point is the limited stability and hence lifetime of such structures, and the availability of, e.g., rechargeable MnO_2-Li cells, in which the (bismuth doped) MnO_2 is also a relatively cheap material.

The use of the high internal area of electronically conducting polymers should make such materials good (cheap) electrochemical capacitors (see Section 13.19) where there would be much to gain (economically!) over, say, porous RuO_2. In condenser discharge, no net current flows so that ohmic losses in pores are not relevant.

TABLE 11.3
Carrier Densities and Mobilities

Polypyrrole Counter-ion	Conductivity $(S\ cm^{-1})$	Carrier Density (cm^{-3})	Carrier Mobility $(cm^2\ V^{-1}\ s^{-1})$
p-Toluene sulfonate	60 ± 2	$3.5 \pm 0.5 \times 10^{20}$	1.1 ± 0.2
Perchlorate	17 ± 4	$3 \pm 2 \times 10^{20}$	0.7 ± 0.5
Tetrafluoroborate	1.3 ± 0.6	$0.14 \pm 0.8 \times 10^{20}$	0.9 ± 0.8

Fig. 11.13. Cyclic voltammograms of 5 m*M* quinone-hydroquinone and polypyrrole-*p*-toluene sulfonate. (a) pH 1, (b) pH 3, and (c) pH 5. (Reprinted from D. L. Miller and J. O'M. Bockris, "Structure of the Polypyrrole/Solution Interphase," *J. Electrochem. Soc.* **139**:970–975, 1992. Reproduced by permission of The Electrochemical Society, Inc.)


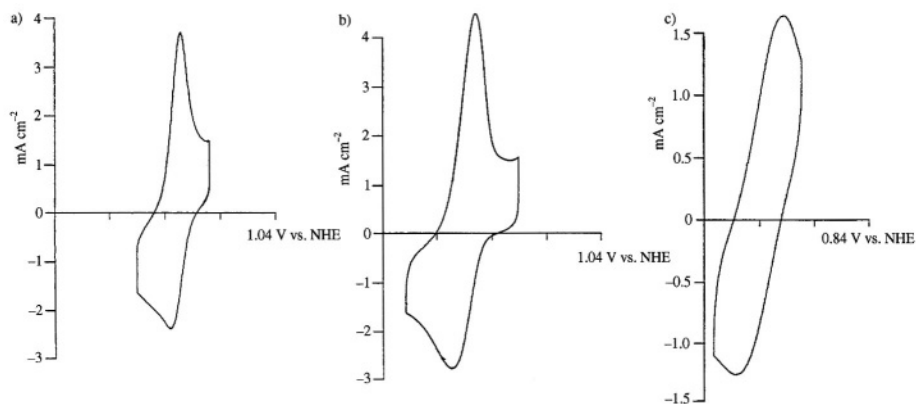**Fig. 11.14.** Tafel plots of quinone–hydroquinone at polypyrrole-*p*-toluenesulfonate: (a) pH 1, (b) pH 3, and (c) pH 5. (Reprinted from D. L. Miller and J. O'M. Bockris, "Structure of the Polypyrrole/Solution Interphase," *J. Electrochem. Soc.* **139**:970–975, 1992. Reproduced by permission of The Electrochemical Society, Inc.)
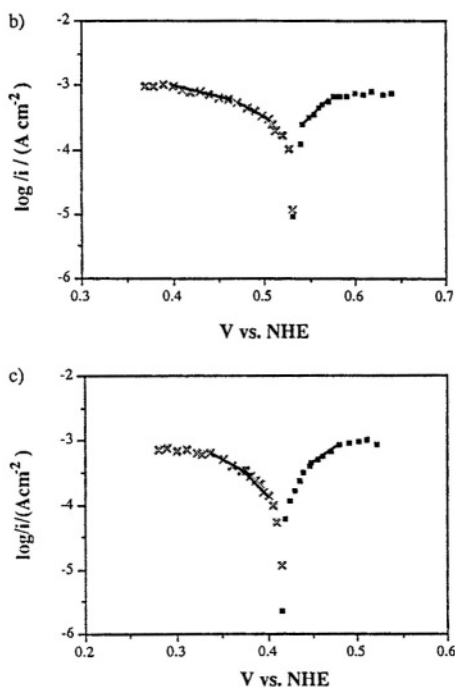
Fig. 11.14. Continued

Other possible electrochemical uses of electronically conducting polymers are discussed in Section 4.9.2 of Vol. 1.

11.5.3.4. *Electronically Conducting Organic Compounds: Problems and the Future.* There is clearly a future in the use of conducting organic compounds as electrodes in many areas of electrochemistry (Section 4.9.2), yet two large areas demand attention in fundamental electro-organic laboratories before much industrial progress can be made.

1. *Predictability of organic structures susceptible to successful ionic doping.* The field is as yet too full of an empirical outlook. Until a convincing and agreedupon model is available for the mechanisms of ionic doping, it may be difficult to predict, for example, the effect of changes in the structure of the organic compounds upon the conductivity that the ionic doping brings to some.

2. *Surface studies of electronically conducting organic compounds are as yet relatively rare.* The structure of the semiconductor/solution interface of these polymers in the negative potential range is open to spectroscopic methods of investigation, as shown by Schultze et al. (1993). However, it is the interfacial region on the anodic side where (for a specific potential region) the compounds are highly conducting, and this needs more study.

$$C_{SC}=20\mu Fcm^{-2} \quad C_{DL}=37\mu Fcm^{-2}$$
$$C_{SS}=0.4\mu Fcm^{-2}$$

V_{FB} =0.23V vs. NHE

"Good" surface state

Valence Band

Polypyrrole

Electrolyte

K^{-1}=15Å
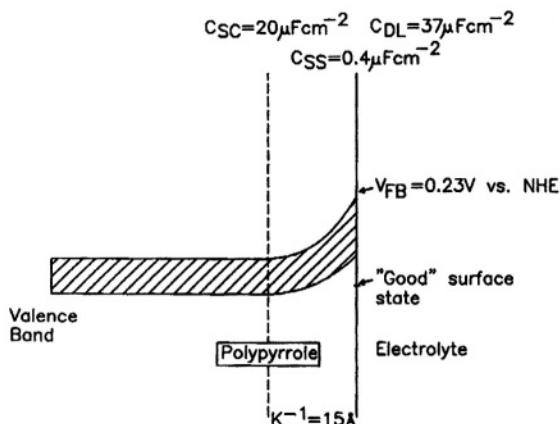
Fig. 11.15. Semiconductor model of the interphase for polypyrrole–aqueous solution. (Reprinted from D. L. Miller and J. O'M. Bockris, "Structure of the Polypyrrole/Solution Interphase," *J. Electrochem. Soc.* **139**:970–975, 1992. Reproduced by permission of The Electrochemical Society, Inc.)
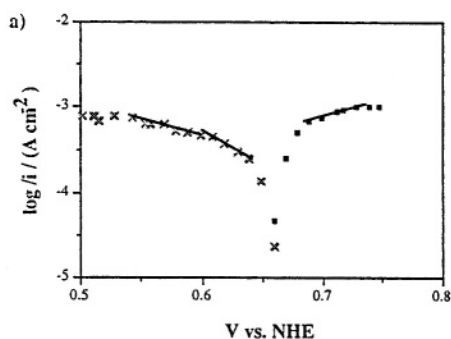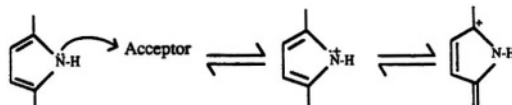

Fig. 11.16. Mechanism for charge transfer through the nitrogen atoms. (Reprinted from D. L. Miller and J. O'M. Bockris, "Structure of the Polypyrrole/Solution Interphase," *J. Electrochem. Soc.* **139**:970–975, 1992. Reproduced by permission of The Electrochemical Society, Inc.)
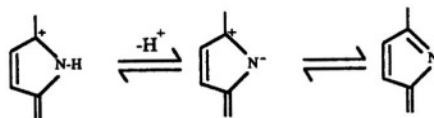
Fig. 11.17. Loss of electrocatalytic activity by deprotonation of the nitrogen within the pyrrole ring. (Reprinted from D. L. Miller and J. O'M. Bockris, "Structure of the Polypyrrole/Solution Interphase," *J. Electrochem. Soc.* **139**:970–975, 1992. Reproduced by permission of The Electrochemical Society, Inc.)
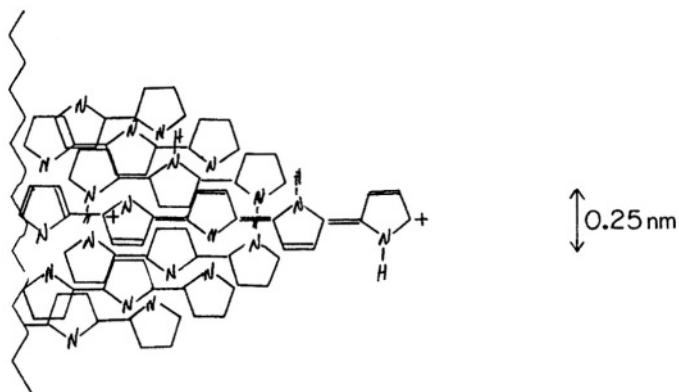
Fig. 11.18. Schematic of a possible microelectrode tip employing polypyrrole. (Reprinted from D. L. Miller thesis, Texas A&M University, 1992.)
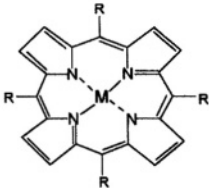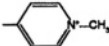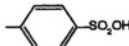
R	Name	
	M = H₂	M = Fe
	a. H₂(2-TMPyP)	f. Fe(2-TMPyP)
	b. H₂(3-TMPyP)	g. Fe(3-TMPyP)
	c. H₂(4-TMPyP)	c. Fe(4-TMPyP)
	d. H₂(2-TAPP)	i. Fe(2-TAPP)
	e. H₂(4-TSPP)	j. Fe(4-TSPP)

Fig. 11.19. Structures of porphyrins and iron porphyrins. (Reprinted from Y. O. Su, T. Kuwana, and S. M. Chen, *J. Electroanal. Chem.* **228**, copyright 1990, p. 177, with permission from Elsevier Science.)

11.6. DESIGNER ELECTRODES

11.6.1. Introduction

In Chapter 7, the electrocatalytic power of various metal surfaces was discussed (Section 7.11). However, as early as 1947, it had been suggested by Heyrovsky that an increase in the rate constant for some electrode reactions could be obtained if it was mediated through another molecule already adsorbed on the electrode surface. The idea was little used until the 1960s, the heyday of attempts to burn organic fuels in fuel cells. Then, Smith (1964) used a number of what he called chelate electrocatalysts (Table 11.4), complex organic structures usually adsorbed on carbon disk electrodes, and found that their presence did indeed cause an increase in the rate at which, e.g., formic acid could be electro-oxidized. Among the compounds used by Smith were the porphins, tetrapyrrole heterocycles, and porphyrins, which are porphins, but with a central metal atom (Fe, Co, Ni) to which the organic compound coordinates through its nitrogen groups (see Fig. 11.19).

A large amount of attention has been given to porphyrins adsorbed on carbon electrodes as catalysts for O_2 reduction (Coleman and Anson, 1980s). The idea of work with these compounds was to fine tune their structure to form, for example, a cage for O_2, stretching this molecule appropriately between two metal atoms in the porphyrin, thus weakening the O–O bond and increasing the rate of reduction of the molecule. The molecular structure of the organics as the factor dominating the catalysis of the reduction of O_2 seems, however, to be inconsistent with the fact (Bagotsky and Tarasevitch, 1977) that heating such catalysts in helium to 800 °C (thus destroying their structures) actually improves the catalysis and the stability of whatever is left on the surface after this severe thermal treatment.[9]

The suggestion of a catalytic effect of adsorbed organics on electrodes has led R. Murray (in particular) to see a field of "the molecular design of electrode surfaces." The idea that organic (or Si-containing) structures on metal surfaces can in effect form new (intelligently designed) electrode surfaces has been encouraged by the possibilities offered by the surfaces of electronically conducting polymers that clearly could be modified in a way that would achieve catalysis. The field was still a young one in the 1990s. It has limitations because of the possible oxidation or reduction of the organic coating. However, metals, too, are oxidized at certain anodic potentials and evolve hydrogen in aqueous solution if pushed too far in the cathodic direction. They have a window of stability and so do organic surfaces designed to promote catalysis.

---

[9]What is left on the electrode surface after the heat treatment consists largely of the metal atoms of the porphyrin. The distances between such metal atoms will correspond to the diameter of the adsorbed organic. Correspondingly, the high temperature will have bound the atom to the carbon surface. Thus, the individual atoms would be acting catalytically; they would not be able to lose catalytic power by aggregating to form clusters in which most of the atoms are hidden from reactants inside the clusters.

TABLE 11.4
Some Ligands Used in Preparation of Chelate
Electrocatalysts

Ligand	Structure	Relative field strength
1. Hexafluoroacetylacetone (HFAA)		Weak
2. Acetylacetone (AA)		Weak
3. Bisacetylacetoneethyle-nediimine (BAAEDI)		Moderate
4. Bissalicylaldehydeethyle-nediimine (BSAEDI)		Moderate
5. Tetraphenylporphin (TPP)		Strong
6. Tetrabenzodiazoporphin (TBDP)		Strong
7. Phthalocyanine (PC)		Strong

11.6.2. Formation of Monolayers of Organic Molecules on Electrodes

The reversible adsorption of organic molecules on electrodes was been discussed in (Section 6.9.3). Here, one is interested in irreversible adsorption, organic molecules that form a film and stick on the electrode through washings and also when the new electrode is used in various reactions.

There are several ways of detecting and examining such films. The first is by observing a characteristic shape in a cyclic voltamogram (Fig. 11.20). Such voltammograms can be interpreted (Chapter 8) with a number of simplifying assumptions. For example, the separation of the peaks in the two directions can yield a rate constant for whatever reaction is occurring at the organic film/solution interface. Much is to be "judged" from the shape of the waves. Potentiostatic transients, ac impedance, and rotating disk are all methods that can be used to examine the films, together with the spectroscopic methods discussed in Chapter 7.

Little detail is known about the irreversible bonding of the organic to the metal or carbon substrates. When aromatics are involved, π bonding is the dominating effect (Blomgren and Jesch, 1961). If the metal on which the organics are irreversibly adsorbed has a d character, exchange of electrons between the d band and the organic molecule may occur.

The hydrophobic layer, in which one end of the adsorbing molecule involves a water-repellent lengthy hydrocarbon chain and the other an anionic (hence hydrophilic) group, is made in two ways. In the one (Section 14.5.4), a Langmuir–Blodgett trough is used to push together the long-chain molecules; when they are in contact, they are lifted off, out of the solution onto, say, a glass slide covered with an evaporated gold film. An example would be

$$[CH_3-(CH_2)_{15}-NO_3]_2$$
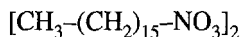
transferred to a conducting tin oxide-covered glass slide.

The other way such layers are formed is by self-assembly. This would involve, say, fatty silanes (SiH_4 substituted with long-chain fatty acid groups) self-assembled onto silica. The thickness of these layers can be varied and the rate of electron tunneling through them to redox ions in solution measured to test the well-known Gamow equation (Section 9.4.3) (Abruña, 1992). Bonding occurs to that metal on the oxide surface, even though the entity is an oxide (Ottagawa, 1984). On carbon, bonding to the basal plane has been identified (Fig. 11.21).

Surfaces of carbon often have on them COOH groups that are difficult to remove. It is only if such surfaces are "activated" by the use of $SOCl_2$ or CH_3COCl that they can be used for adsorption and the formation of a new surface (see Fig. 11.22).
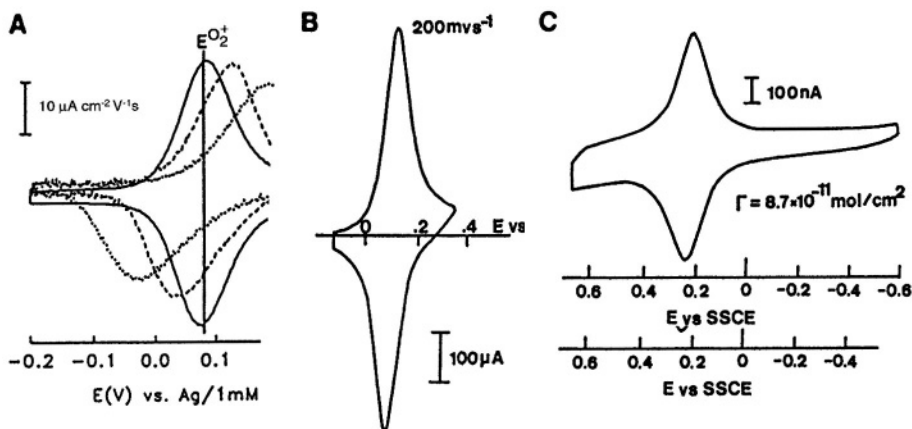
Fig. 11.20. Illustrative cyclic voltammograms of electrode surface-immobilized redox molecules. (A) A mixed monolayer of $(\eta^5 C_5H_5)Fe$ $(\eta^5 C_5H_4)CO_2(CH_2)_{16}SH$ and $CH_3(CH_2)_{15}$ SH self-assembled on Au, in 1 M $HClO_4$. Potential scan rate: 10 mV s^{-1} (solid line), 120 mV s^{-1} (dashed line), 500 mV s^{-1} (dotted line), an electron-transfer distance effect (Reprinted from *Techniques of Chemistry, Molecular Design of Electrode Surfaces,* R. W. Murray, ed., Vol. 22, p. 18. Copyright © 1992 John Wiley & Sons. Reprinted by permission of John Wiley & Sons, Inc.) (B) Multilayer film (5×10^{-9} mol cm^{-2}) formed on Pt by hydrolysis of N,N,N',N'-tetrakis(trimethoxysilyl-3-propyl)-benzenediamine in 0.1 M Bu_4NClO_4/CH_3CN (Reprinted from *Techniques of Chemistry, Molecular Design of Electrode Surfaces,* R. W. Murray, ed., Vol. 22, p. 15. Copyright © 1992 John Wiley & Sons. Reprinted by permission of John Wiley & Sons, Inc.) (C) Self-assembled monolayer of $[Os(bpy)_2 \times (dipy)Cl]^{1+}$, where dipy = 4,4'-trimethylenedipyridine, on Pt, in 0.1 M $Bu_4NCl)_4/CH_2Cl_2$, 100 mV s^{-1} (Reprinted from *Techniques of Chemistry, Molecular Design of Electrode Surfaces,* R. W. Murray, ed., Vol. 22, p. 12. Copyright © 1992 John Wiley & Sons. Reprinted by permission of John Wiley & Sons, Inc.)

11.6.3. Apparent Catalysis by Redox Couples Introduced into Polymers Attached to Electrodes

A number of electrode reactions have been carried out on carbon surfaces containing preadsorbed redox polymers of varying thicknesses (thus designing a new electrode surface). In respect to the oxidation of Fe^{2+}, a protonated vinyl film has been used containing $(IrCl_6)^{3-/4-}$ as the redox couple. The reaction rate at a given overpotential per *geometric* unit area is increased. There is dual control of the reaction rate between interfacial electron transfer and diffusion of the Fe^{2+} reactant into the film.

More interesting because of earlier work on the reduction of O_2 at macrocycles is the reduction of O_2 in dimethyl sulfoxide using a redox couple involving anthraquinone mixed into various complex polymers that are attached to the underlying metal or graphite. The diffusion of O_2 into the film turns out to be relatively fast, as
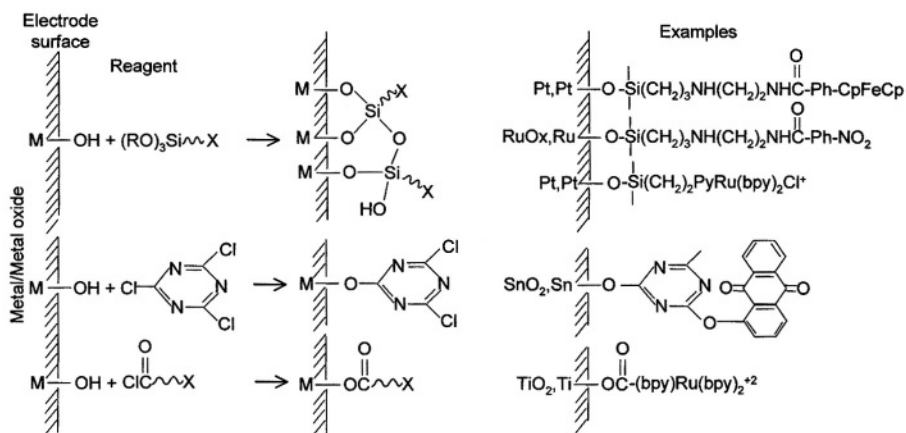
Fig. 11.21. Example structures for monolayer covalent bonding to metal oxide electrodes. (Reprinted from *Techniques of Chemistry, Molecular Design of Electrode Surfaces,* R. W. Murray, ed., Vol. 22, p. 111. Copyright © 1992 John Wiley & Sons. Reprinted by permission of John Wiley & Sons, Inc.)

determined by a rotating disk electrode, which also defines the film thickness at the polymer/solution interface. The rate-determining step depends on the film thickness; at low thickness it is an activated electron-transfer step and at higher thicknesses it is diffusion within the film coupled to interfacial electron transfer.

There are examples of designer electrodes in which two materials are necessary. For example, $(Ru(NH_3)_6)^{3+/2+}$ serves to transport electrons by means of a hopping
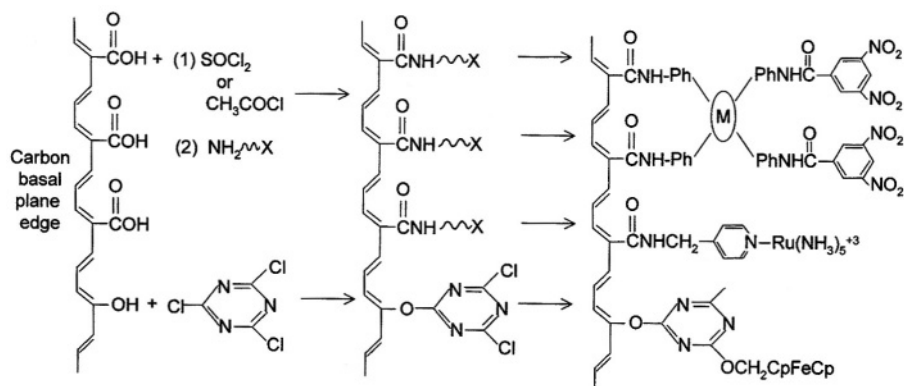

Fig. 11.22. Example structures for monolayer covalent bonding to edge plane pyrolytic graphite electrodes. (Reprinted from *Techniques of Chemistry, Molecular Design of Electrode Surfaces,* R. W. Murray, ed., Vol. 22, p. 143. Copyright © 1992 John Wiley & Sons. Reprinted by permission of John Wiley & Sons, Inc.)

mechanism. Catalysis is served by cobalt III/II tetrakis (4-*N*-methyl pyridyl) por-phyrin. The latter substance is not a good electron carrier, but together they reduce O_2 to H_2O_2.

11.6.4. Conclusion

Many expertly designed new electrode surfaces have been created (R. Murray, 1992–1998) and made active by mixing the polymer with some kind of redox system, which in turn brings about the oxidation or reduction of the reactant. Several factors are intertwined here: the permeation of the reactant into the film which reaches an optimal thickness, electron-transfer steps, and steps involving surface reaction.

There are undoubtedly systems where the net result of the use of macrocycles is advantageous (e.g., use of the crown ethers by Wass in 1989 to photo reduce CO_2); it is necessary to be sure that the increase in rate is more than an effect of the extra area introduced (which is somewhat similar to that arising from the use of a porous electrode containing a metal catalyst) before one concludes that there is a catalytic effect per real square centimeter due to the new structure.

Further Reading

Seminal

1. A. Von Arnim, *Ann. Phys. (Leipzig)* **8**:257 (1801). Electrolysis of red wine.
2. A. Crum-Brown and J. Walker, *Justus Liebigs Ann. Chem.* **261**:107 (1891). Kolbe synthesis.
3. H. Hofer and M. Moist, *Justus Liebigs Ann. Chem.* **323**:285 (1902). Alcohols from carbonylic acid.
4. R. Criegee, *Justus Liebigs Ann. Chem.* **73**:523 (1936). First chiral synthesis.
5. A. Hickling, *Trans. Faraday Soc.* **38**:27 (1942). The potentiostat.
6. J. R. Backhurst, J. M. Coulion, F. Goodridge, R. E. Mimley, and M. Fleischmann, *J. Electrochem. Soc.* **116**:1600 (1969). Packed-bed (3D) electrodes.

Modern

1. J. Heinze, K. Hinkelmann, M. Dietrich, and J. Mortenzen, *Ber. Bunsenges. Phys. Chem.* **89**:1225 (1985). Synthesis of conducting organic polymers.
2. A. V. Rama Rao, M. K. Gurjar, and S. V. Jocki, *Tetrahedron Asym.* **1**: 697 (1990). Formation of chiral compounds.
3. B. R. Scharifker, E. Garcia-Partorize, and W. Marino, *J. Electroanal. Chem.* **300**: 85 (1991). Growth of polypyrrole films.
4. D. Stöckert, R. Kessel, and J. W. Schultze, *Synthet. Metals* **41–43**: 1295 (1991). Photoelec-trochemistry of heterocyclic polymers.
5. D. L. Miller and J. O'M. Bockris, *J. Electrochem. Soc.* **139**: 967 (1992). The polypyr-role/solution interface.
6. T. Yoshiyama and T. Fuchigami, *Chem. Lett.* **40**: 1995 (1992). Difluorination of diacetals.

7. D. J. Fermin, J. Mostany, and B. R. Scharifker, *Curr. Topics Electrochem.* **2**:132 (1993). Review of electronically conducting organic compounds.

8. "Electrons are green:Electrochemical Syntheses," RC-175. Business Communications Co., Inc., Niagara Falls, NY (1994).

9. E. Herrero, K. Frenasczuk, and A. Wieckowski, *J. Phys. Chem.* **8**: 5074 (1994). Methanol eoxidation at low-index crystal planes.

10. N. L. Weinberg and D. J. Mazur, *Annual Report of the Royal Society of Chemistry* **42**: 58 (1996). Recent electro-organic syntheses.

11. R. D. Little, H. Bode, T. Brigant, and M. Schwaebe, in *Novel Trends in Synthesis,* S. Torii, ed., p. 103, Kodanski, Tokyo (1995). Stereoselective coupling.

12. J. Y. Lin, W. K. Paik, and I. H. Yeu, *Synth. Met.* **69**: 451 (1995). Growth of polypyrrole and the microbalance.

13. M. Palucka, G. J. McCormack, and E. N. Jacobson, *Tetrahedron Lett.* **36**: 547 (1995). Asymmetric epoxidation of olefin gives a chiral result.

14. J. Um and W. K. Paik, *Bull. Korean. Chem. Soc.* **17**: 369 (1996). Capacitance of polypyrrole films.

15. D. J. Fermin, Helene Teruel, and B. R. Scharifker, *J. Electroanal. Chem.* **401**: 207 (1996). Species and charge carriers in the oxidation of polypyrrole.

16. D. Kim, D. Lee, and W. K. Paik, *Bull. Korean. Chem. Soc.* **17**: 707 (1996). Polypyrrole films and ellipsometry.

17. C. Herrero, W. Chrzanowski, and A. Wieckowski, *J. Phys. Chem.* **99**: 10423 (1995). Dual-path mechanism in the electro-oxidation of methanol.

18. Y. Sato, M. Fujita, F. Mizutani, and K. Uosaki, *J. Electroanal. Chem.* **409**: 145 (1996). Electrochemical properties of 2-mercaptohydroquinone monolayers on gold.

19. T. Kondo, T. Itu, S. Nomura, and K. Uosaki, *Thin Solid Films* **286**: 652 (1996). Photoelectrochemical properties of porphyrin mercaptoquinone.

20. S. Ye, A. Yashiro, Y. Sato, and K. Uosuki, *J. Chem. Soc. Faraday Trans.* **92**: 3813 (1996). FT-IRCAS studies of self-assembled 2-(11-mercaptoundecyl) hydroquinone.

21. M. E. William, J. Long, H. Mesui, and R. W. Murray, *J. Am. Chem. Soc.* **119**:1286 (1997). Electron transport in redox polyether melts.

22. R. H. Tersil, J. E. Hutchison, and R. W. Murray, *J. Phys. Chem.* **101**: 1535 (1997). Electron hopping in viologen tetraethylene oxide copolymer.

23. V. M. Manes, H. Masui, R. M. Wrighton, and R. W. Murray, *J. Am. Chem. Soc.* **119**: 3987 (1997). Luminescence in molten ruthenium 2,21-bipyridine complex.

EXERCISES

1. The most famous application of electrochemistry to industry in modern times is the synthesis of nylon. (a) What step in the synthesis is electrochemical? (b) Is the synthesis in an *aqueous* solution? (c) What part is played by NEt_4^+? (d) Would the cell concerned have to have a membrane and if so, when? (Bockris)

2. One of the more important discoveries in organic chemistry in the second half of the twnetieth century has been the knowledge that some organic polymers can be made highly (electronically) conductive, (a) Give three examples. (b) What does "doping" mean in respect to those compounds and how does it differ from the "doping" referred to in semiconductors? (c) Why is it that the electronically conducting polymers can only be used within a restricted range of potentials? (Bockris)

3. The following data are for the cathodic reduction of thyroxine (one of the iodoamino acid derivatives) at a silver electrode. A disk electrode was rotated at various angular velocities and the current was measured [see R. A. Osteryoung et al., *Anal. Chem.* **56**:1202–1206 (1984)].

$\omega^{1/2}$	I_d (in μA)
16.18	83
12.94	68
9.71	51
6.47	35

$A = 0.442 \text{ cm}^2$
$v = 0.01 \text{ cm}^2/\text{s}$
$C = 5.0 \times 10^{-8} \text{ mol/cm}^3$
$D = 3.3 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$

(a) Show that the process is diffusion controlled (in the range of $\omega^{1/2}$). (b) Evaluate the number of electrons involved in the reduction. (c) Justify. (Contractor)

4. The redox mediator 2,6-dichlorophenol indophenol, can mediate electron transfer from and to the redox enzyme, cytochrome *c*. The mediator was switched between the oxidized and reduced forms by the application of a potential using optically transparent electrodes in a thin-layer cell. From the absorbances of the oxidized and reduced form of the enzyme, the ratio of their concentrations at various potentials was obtained. Calculate the formal potential $E°$ of the enzyme from the data given in Table E.1. Confirm that the enzyme redox process involves one electron transfer. (Contractor)

TABLE E.1

$E_{applied}$ (mV vs. SCE)	−30	−10	+10	+30	+54
log [o]/[R]	−0.80	−0.53	−0.21	+0.1	+0.5

PROBLEMS

1. In contemplating the wisdom of going down an electrochemical rather than a chemical path in synthesizing organic compounds, the cost of the electricity is

sometimes brought out as a point against an electrochemical route.[10] In chemistry, reactions only proceed down the gradient of the chemical potentials, but in electrochemistry they proceed down the gradient of the electrochemical potentials and hence one can mount *up* a free energy gradient. In such a case, there is no chemical alternative (cf. the electrosynthesis of hydrogen from water). However, what has been somewhat neglected are the possibilities of electrochemical synthesis where $\Delta G_{chemical}$ is negative and the reactions can be carried out in a fuel cell, producing commercially valuable electricity as the by-product of the synthesis of the valuable product.

Consider the synthesis of the following substances: (a) Dichlor ethylene (from C_2H_4 and Cl_2), (b) benzaldehyde (from toluene and O_2), and (c) Adiponitrite (from acrylonitrile by dimerization).

(a) Write out the overall reaction in each case. (b) Break this reaction down into two electrode reactions. Using tables of thermodynamic data, calculate on the hydrogen scale, the standard free energy of each electrode reaction and hence of the overall reaction. If it is negative in the direction of synthesis, there is a possibility of a fuel cell process. (c) Determine whether this is so for each of the above reactions. (Further consideration, e.g., of the electrocatalysis is outside the scope of this problem.) (Bockris)

2. Consider the following information on the electrochemical oxidation of methanol on platinum: It is known that upon adsorption of CH_3OH upon platinum, dehydrogenation occurs. The final product is CO_2. Several analyses made on the basis of potentiodynamic sweep data suggest the presence on the electrode surface of $\check{c}$ −OH. However, further spectroscopic work, particularly that in potentiostatic work in steady state finds linear CO, i.e., $\check{s}$=O, the intermediate. The Tafel slope is 60 mV at low current densities and 120 mV at higher values of the current. Potentiodynamic profiles show that the overall number of electrons in the oxidation is 1.2–1.5 (i.e., two kinds of CO radicals of somewhat different character as to their adsorption are on the surface).

(a) On the basis of the information given, write a series of steps that go from CH_3OH in solution to CO_2 in the gas phase. (b) Show that your mechanism fits the data. (c) Work out an equation for the *i–V* relation. (d) How does it fit the results stated above? (Bockris)

3. The electrode material is economically important in electrochemical synthesis and an attempt is made to reduce to a minimum the use of expensive noble metals, which are regularly used in small quantities in academic investigations. Of course, the principal danger in using less stable materials is that they themselves will be oxidized or embrittled in use, respectively, as anode or cathode.

---

[10]According to the experienced German organo-electrochemist, Fritz Beck, the cost of the electricity in electrochemical synthesis is seldom more than 10% of the whole.

Among conventional electrode materials are Pb and Hg as cathodes (they have a high H overpotential and hence are likely to prefer to give electrons to compounds that will accept them at less cathodic potentials than that needed for H_2 solutions). PbO_2, Fe_2O_3, and SiC are examples of useful anode materials. Since about 1980, several newer electrode materials have found a place. For cathodes, Ni has become important. Cr_2O_3, TiC, and carbon-black-filled polymers have been used as anodes.

Suggest cathodes and anodes for the following processes, giving a full rationalization of your choices: (a) methyl chloride to tetraalkyl lead, (b) acetone to pinacol, and (c) nitrobenzene to *p*-amino phenol. Your choices need not be limited to the electrode materials mentioned here. (Bockris)

4. In the reduced state, poly-3-methyl thiophene is an intrinsic semiconductor with an energy gap of 1.95 eV. (a) What would be the value in electron volts of its Fermi level? (b) Show that the concentration of electrons in the conduction band would be negligible. However, from 0.45 to 1.1V on the normal hydrogen scale, this polymer shows metallic conductivity. (c) What change in electronic structure has occurred?

Among the several fields in which electronically conducting polymers are useful or may be so the future, are electrocatalysis, prosthetics, and electrodes suited for use with biomaterials. (d) Consider each of these areas and state the reasons you think electronically conducting compounds (those now available and those that may be synthesized) would have characteristic properties of special use in the areas mentioned. (Bockris)

5. A method for the preparation of aniline consists in the reduction of nitrobenzene on an electrode of a platinum amalgam (cathode) in acidic media and inert atmosphere. The reversible potential for the system nitrobenzene–aniline is, in this media, 0.87 V (vs. normal hydrogen electrode at 25°C). The overpotential for the hydrogen evolution reaction on the cathode is very large, and the electroreduction reaction of nitrobenzene has a $j_0 = 10^{-8}$ A cm^{-2} and a Tafel slope of -0.120 V dec^{-1}. (a) Calculate the rate of aniline production in mol cm^{-2} s^{-1} if the cathodic potential that the system nitrobenzene–aniline develops is -0.13 V vs. the same reference. Consider the mass transfer phenomena negligible. (b) If the process takes place in a reactor of two compartments, calculate the current density for the aniline production if the applied potential is 1.0 V. What parameters could be modified to obtain a maximum current density? Consider that the anodic process is the electro-oxidation of water with a reversible potential of 1.2 V. Assume that the total electrical resistance for both cell compartments and the union between them is 10^4 Ω. The Tafel parameters for the electro-oxidation of water are $j_0 = 10^{-10}$ A cm^{-2} and $(\partial\eta/\partial \log j)_T = +0.12$ V dec^{-1}. Consider that in the cathode, 10^9 particles of mercury of 0.1-μm are distributed on the platinum metal. (c) Discuss what phenomena might occur as a consequence of an incomplete sealing in the cathodic compartment. (Zinola)

This page intentionally left blank

CHAPTER 12

ELECTROCHEMISTRY IN MATERIALS SCIENCE

12.1. CHARGE TRANSFER, SURFACE, AND CIVILIZATION

12.1.1. Introduction

In the early stages of human civilization, humans preserved themselves in a hostile environment by functioning as bioelectrochemical machines, converting the solar energy stored in food via electrochemical reactions into muscle power. But recently man has become increasingly a *cyborg,* a person of artificial parts and organs; he has linked himself more and more with machines that harness nuclear energy and the solar energy stored in coal and oil and has in this way satisfied his needs with increasing efficiency. Thus, the progress of civilization has been marked by an increasing use of machine power and a decreasing use of muscle power. But, what of the future? The trend is clearly that man will make minimal use of his own biochemical energy converters and turn to machines to effect the conversion of energy into convenient forms. Man is bound to lean increasingly on computerized mechanisms to do the work that he wants.

To make this vision a reality, the machines that do man's bidding and thus become the basis of the material aspects of civilization must be able to function without decay over years in the terrestrial atmosphere. Materials, mainly metals, used in fabrication must be stable. If the metals become unstable, then machines fabricated partly from these metals undergo an undesired obsolescence. An industrial civilization depends in a crucial way upon the stability of metals in its moist (and often impurity-containing) atmospheres.

It is an interesting fact that a piece of metal remains stable for an almost indefinite period of time if it is stored in a vacuum. It appears that metals acquire stability when their surfaces are isolated from the normal terrestrial environment. If this isolation is

not achieved, metals become unstable in various ways. They develop cracks and break upon strain with catastrophic suddenness. They suffer fatigue, i.e., loss of strength, when subjected to periodic stress. They undergo a process of embrittlement. Their surfaces are transformed into oxides that peel off, or they just dissolve away. With the exception of the (hence) expensive noble metals, all metals are unstable to varying degrees in a terrestrial atmosphere. The most widely used metals, namely, iron, aluminum, copper, nickel, and alloys of these metals, all decay and lose good mechanical properties in unprotected contact with air.

One conclusion is obvious. The stability of metals is determined by the events at the interface between these metals and their environment. The internal strength of a metal (particularly a metal under stress) is influenced in the long run by events at its surface. If the surface of a metal is stable, its interior tends to remain so. The detrimental transformation of the bulk properties of a metal begins at its surface. This, then, is a link between civilization, surfaces, and, as will be seen, electrochemical reactions.

An important feature of the terrestrial environment must now be noted. The atmosphere is essentially moist air containing dissolved carbon dioxide. (Marine atmospheres consist of moist air often containing sodium chloride in suspension.) Moisture in contact with the terrestrial atmosphere becomes an ionically conducting medium, an electrolyte.

Since metals become unstable (undergo the events named above) when they come into contact with the moist atmosphere, it is reasonable to conclude that this instability of metals results from charge-transfer reactions at their interfaces. This is why the rate of corrosive destruction of a metal's surface is greatly reduced by removal of moisture from the atmosphere. Keeping a metal in a vacuum is equivalent to removal of the electrolyte in contact with the metal and therefore to the prevention of charge-transfer reactions. Thus, the spontaneous instability (or corrosion) of metals results from the charge-transfer reactions at the electrified interface between the metal and the moist, CO_2^- or NaCl-containing air (Wolaston, de la Rive).

Such a view was confirmed by many detailed experiments in the first half of this century. These experiments included direct studies of the rate and products of the corrosion of a metal as a function of the electrolytic conduction of the moist film. They also involved an imaginative extrapolation from experiments on energy-producing electrochemical cells, in which, e.g., separate pieces of zinc and copper were immersed in an electrolytic solution, to a situation in which an actual piece of impure (copper-containing) zinc decayed when brought into contact with a film of moisture that contained dissolved electrolyte.

12.1.2. A Corroding Metal Is Analogous to a Short-Circuited Energy-Producing Cell

Charge-transfer reactions are the basis of electrochemical substance-producing cells driven by an external current source and of electrochemical energy-producing cells driving an external load (see Chapter 13). Metallic corrosion, it has been stressed,

also arises from the electrodic charge-transfer reactions at the interface between a metal and its electrolytic environment. But, in the case of a corroding metal, where is the external source driving the charge-transfer reactions, or where is the external load consuming the current produced by the charge-transfer reactions? The conceptual relationship between electrochemical cells and corroding metals must be developed.

Suppose that a piece of zinc and a piece of copper are immersed in an electrolyte containing Zn^{2+} and Cu^{2+} ions (Fig. 12.1). It has been argued that, because the equilibrium potential of the $Zn^{2+} + 2e \rightleftharpoons Zn$ reaction is negative with respect to that for the $Cu^{2+} + 2e \rightleftharpoons Cu$ reaction, the zinc electrode is negatively charged with respect to the copper electrode. When an external electron path is provided by connecting the zinc and copper electrodes through an *external* load of resistance R_e, electrons flow through this external circuit from the zinc to the copper. To keep this current going, the zinc electrode dissolves to form Zn^{2+} ions, and the copper deposits on the copper electrode. Hence, a Zn–Cu *electrode couple* acts as an energy producer. The potential difference across such cells has been analyzed (see Section 7.13.6), and it has been shown that the potential difference decreases with the cell current (Fig. 12.2). But this cell current is decided by the external load; make its resistance R_e lower, and the cell current increases.
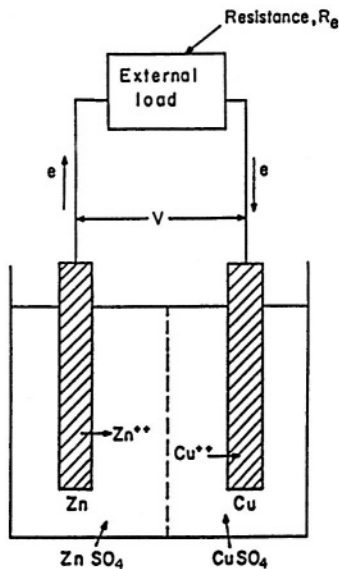

Fig. 12.1. In a Daniel cell, the current is kept going by the dissolving zinc and depositing copper ions.
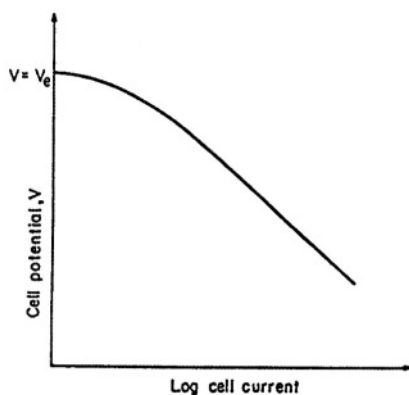
Fig. 12.2. The cell potential decreases as increasing current passes from the cell through the load.

What happens when the external resistance is made zero, i.e., when the copper and zinc electrodes are brought into electrical contact, or short-circuited (Fig. 12.3). Of course, the copper continues to deposit and the zinc continues to dissolve at a certain current, but the potential difference across the cell will become zero. This thought experiment is equivalent to what happens when a bar of copper and a bar of zinc are welded together and put into an electrolyte containing cupric ions (Fig. 12.4). The zinc dissolves as the copper deposits. Similarly, if, e.g., iron is welded together with some other metal and placed in an electrolytic solution, whether it dissolves will depend on whether its equilibrium potential is more negative or more positive than that of the other metal.

The next step in the thought experiment consists in taking a large number of strips of copper and zinc and joining them so that there are alternate strips of the two metals, a multiband arrangement. If this assembly is immersed in solution containing cupric ions, the copper strips will be the sites for copper deposition and the zinc strips will be the sites for zinc deelectronation. Once again, the net result is copper deposition and zinc dissolution.

Finally, think of a bar of zinc with microscopic inclusions of copper, i.e., with copper impurities, (Fig. 12.5). If this zinc bar is immersed in a solution containing Zn^{2+} and Cu^{2+} ions, the result is that the zinc will dissolve out and copper will deposit preferentially on the already existing copper areas or even form such areas by crystallization.

But notice that in all these thought experiments to keep the zinc dissolving, it is not essential that the deposition of copper should be the electronation reaction. Even if the aqueous electrolyte has no Cu^{2+} ions but consists of an ionically conducting moist film, other electronation reactions are possible, e.g., hydrogen evolution, $2H_3O^+ + 2e \rightleftharpoons H_2 + 2H_2O$, or oxygen reduction, $O_2 + 4H^+ + 4e \rightleftharpoons 2H_2O$,
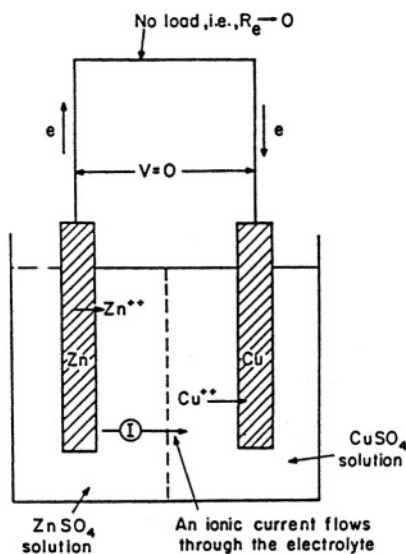
Fig. 12.3. The maximum current given by the cell is attained when the load is removed and the two electrodes are short-circuited.

and, as long as these electronation reactions take place, zinc dissolution will continue (Fig. 12.6). The bar of zinc is undergoing corrosion; it becomes unstable and eventually destroys itself as a consequence of the electrochemical reactions occurring at the interface of the metal and ionically conducting moist films or actual solutions.

Starting from the familiar Zn–Cu cell, the above discussion has shown that by short-circuiting the cell and altering the spatial location of the electron source and sink,

Fig. 12.4. Electrodes are short-circuited also when two welded bars, one of zinc and one of copper, are immersed in an electrolyte containing cupric ions.

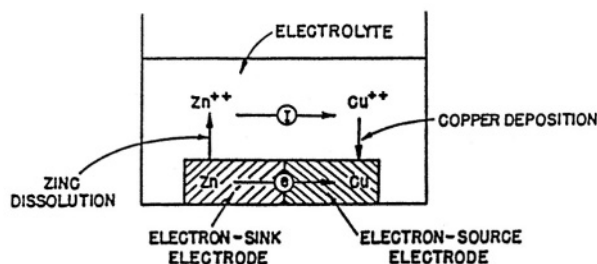one is able to understand the corrosion of a piece of zinc. A corroding metal consists of an electron-sink area at which a deelectronation reaction (i.e., metal dissolution) occurs, an electronic conductor to carry the electrons to the electron-source area where an electronation reaction occurs, and an ionic conductor to keep the ion current flowing and to function as a medium for the electrodic reaction (Fig. 12.7). This model of corrosion is often termed the *local-cell theory of corrosion.*

12.1.3. Mechanism of the Corrosion of Ultrapure Metals

On the basis of the local-cell theory, an ultrapure metal without impurity inclusions would be expected to be incorrodible. In general, the purer a metal, the more stable it is in an aqueous environment. But even an ultrapure metal corrodes. Why?

The basic mechanism for the instability of ultrapure metals was suggested by Wagner and Traud in a classic paper in 1938.[1] The essence of their view is that for corrosion to occur, there need not exist *spatially separated* electron-sink and -source areas on the corroding metal. Hence, impurities or other heterogeneities on the surface are not essential for the occurrence of corrosion. The necessary and sufficient condition for corrosion is that the metal dissolution reaction and some electronation reaction proceed simultaneously at the metal/environment interface. For these two processes to take place simultaneously, it is necessary and sufficient that the corrosion potential be more positive than the equilibrium potential of the $M^{n+} + ne \rightleftharpoons M$ reaction and more negative than the equilibrium potential of the electronation (cathodic) reaction $A + ne \rightarrow D$ involving electron acceptors contained in the electrolyte (Fig. 12.8).

Hence, the present view is a unified one. When the electron-sink and -source areas are distinct in space and stable in time, one has the local-cell, or heterogeneous, theory of corrosion [Fig. 12.9(a)]. On the other hand, when the metal-dissolution and electronation reactions occur randomly over the surface with regard to both space and

---

[1]The Wagner and Traud paper of 1938 is the basis for modern understanding of the mechanism of corrosion. The senior author, Carl Wagner—then dean of science at the University of Damstadt, in Germany—became an MIT professor after WWII. Quiet spoken, modest in bearing, and a bachelor, he had difficulty in adjusting to noisy faculty parties at which attendance was expected and alcohol freely consumed. Thus he would arrive half-way though the event, walk right up to a colleague in a group, cough quietly to gain attention, and begin (typically). "In respect to Section 3 of your paper, I cannot agree with equation 4." He was highly respected, but not really socially popular.

Wagner refused lucrative consulting offers from industry, but studied problems posed to him alone. Sometimes this would result in a letter containing calculations and suggestions. In this way, he initiated study of the three-phase boundary in fuel cell electrodes at the Pratt and Whitney company in the 1950s. His idea became the basis (with electrocatalysis) of successful fuel cell design.

Car Wagner was reluctant to accept graduate students and wrote his papers standing at a podium, in a single draft that was the finished paper. He thought mainly on long walks alone on Cape Cod. He lived like a student in single rooms in boarding houses. When his time came, he checked into a hospital in the German academic center, Göttingen. He asked to be left alone—no visitors. Thus passed on of the great contributors to basic electrochemical materials science.
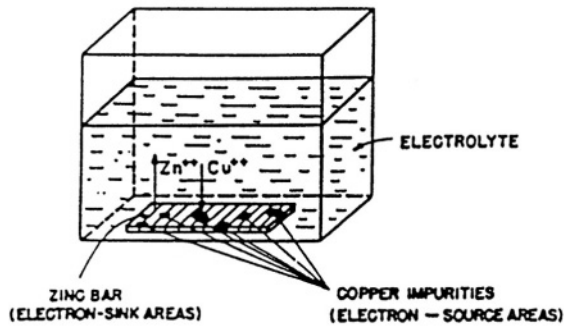
Fig. 12.5. Even without copper strips, copper will deposit on a zinc bar from a cupric ion-containing electrolyte, preferably at some microscopic inclusions of copper impurities.

time, one has the Wagner–Traud, homogeneous theory of corrosion [Fig. 12.9(b)]. The Wagner–Traud mechanism with its random and dynamic deelectronation (anodic) and electronation sites requires a homogeneous metal surface. This is because heterogeneities tend to fix the deelectronation and electronation reactions to stable sink and source areas.

In some practical situations, however, there are heterogeneities of one type or another. Impurities are the most obvious type of heterogeneities, but there are other types, e.g., different phases of an alloy, or a metal with a nonuniform stress distribution or with a nonuniform access to electron acceptors. Thus, the local-cell, or heterogeneous, theory of corrosion has a wide scope of applicability. The homogeneous theory
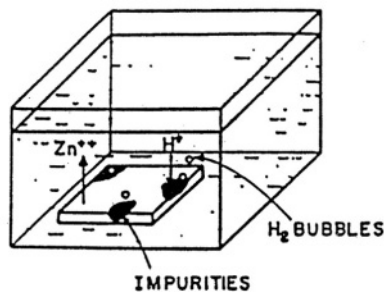
Fig. 12.6. The hydrogen-evolution reaction is the electronation reaction that occurs in zinc corrosion in acid solution.

Fig. 12.7. According to the local-cell theory, corrosion can take place if the piece of metal is in a moist atmosphere. There are separate sites for the deelectronation and electronation reactions.

Fig. 12.8. Corrosion can always occur when the reversible potential of the metal dissolution is more negative than the actual potential of the metal, and that of the electronation reactions is more positive.

Fig. 12.9. A metal can corrode (a) by a heterogeneous mechanism if it contains areas of different electrodic properties or (b) by a homogeneous one if the surface is uniform.

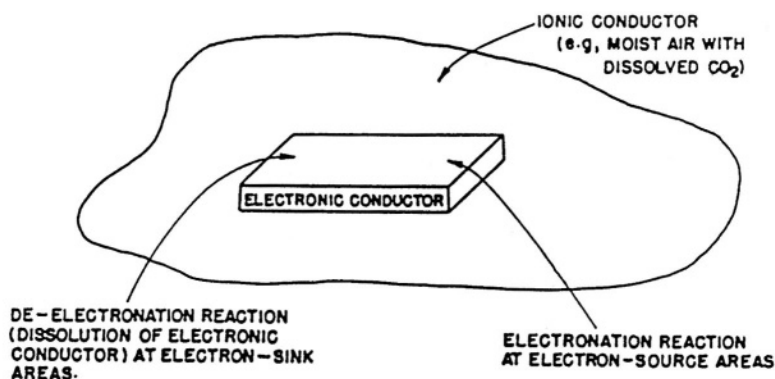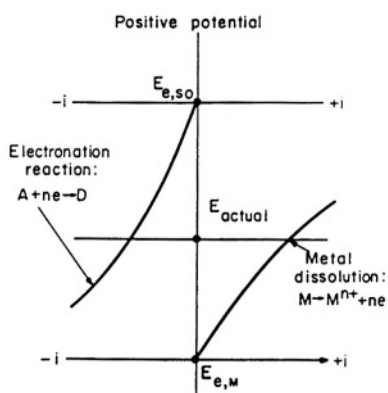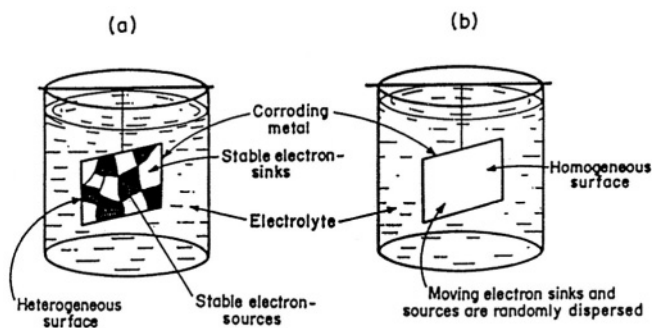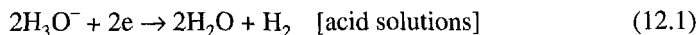of corrosion emphasizes that irrespective of the presence or absence of impurities, *metals become unstable because of different electrodic charge-transfer reactions occurring simultaneously and in opposite directions at the surface.*

12.1.4. What Is the Cathodic Reaction in Corrosion?

A very important aspect of the corrosion of metal has been only touched upon so far. This aspect concerns the electronation (cathodic) reaction required to complete the corrosion circuit by consuming the electrons transferred to the metal from the metal-dissolution reaction. The question is: What is the electronation (cathodic) reaction?

Theoretically, it can be any reaction with an equilibrium potential that is more positive than the equilibrium potential of the metal-dissolution reaction. In practice, it is a reaction of the type of $A + ne \rightleftharpoons D$, where A is an electron-acceptor species present in the electrolyte that is in contact with the corroding metal. In aqueous electrolytes, the electron acceptors invariably present are H_3O^+ ions and dissolved oxygen, the corresponding electronation reactions being

$$2H_3O^- + 2e \rightarrow 2H_2O + H_2 \quad \text{[acid solutions]} \tag{12.1}$$

and

$$O_2 + 4H^+ + 4e \rightarrow 2H_2O \quad \text{[acid solutions]} \tag{12.2}$$

or

$$O_2 + 2H_2O + 4e \rightarrow 4OH^- \quad \text{[alkaline solutions]} \tag{12.3}$$
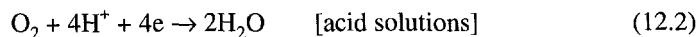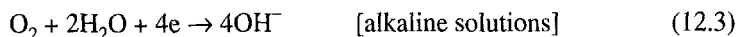
The electrolyte may also contain species such as Fe^{3+} ions or nitric acid, in which case there can be additional electronation reactions of the type

$$\mathrm{Fe^{3+} + e \rightarrow Fe^{2+}} \tag{12.4}$$

or

$$\mathrm{3H^+ + NO_3^- + 2e \rightarrow HNO_2 + H_2O} \tag{12.5}$$

If several electronation (cathodic) reactions are possible, i.e., their equilibrium potentials are positive with respect to the metal-dissolution equilibrium potential, then the one that yields the highest corrosion current is preferentially adopted. There is, of course, no new principle here; when parallel reactions can occur, the current is controlled by that reaction which yields the largest current corresponding to the given potential. This point is brought out c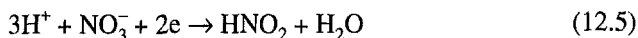learly by the increase in the corrosion rate of iron in oxygenated solution compared with that in a deoxygenated one; the rate in the latter is decided by the $\mathrm{2H_3O^+ + 2e \rightleftharpoons 2H_2O + H_2}$ electronation (cathodic) reaction, whereas in the former, it is determined by oxygen reduction. The higher the pressure of oxygen in the gas phase, the higher the corrosion rate since the solubility of oxygen in the electrolyte is proportional to pressure (Table 12.1). Upon addition of dilute nitric acid, the corrosion rate is increased even more because of the occurrence of reaction (12.5) involving nitrate ions.

12.1.5. Thermodynamics and the Stability of Metals

Suppose that one were faced with the task of deciding whether a particular metal would be suitable as a material of construction or fabrication in a given environment. The problem, e.g., may consist of approving or rejecting the use of a mild-steel reaction vessel in a technologically important process involving an aqueous medium. The real criterion for making the decision on the stability of the iron vessel is the magnitude of the rate of its dissolution; if it has a negligible rate of corrosion and sufficient strength, it is suitable for the purpose.

But suppose that even before one calculates or measures the corrosion rate, one requires a yes or no answer regarding the stability of the steel vessel. The question is: Will the $\mathrm{Fe \rightarrow Fe^{2+} + 2e}$ deelectronation reaction and the electronation reaction, which

TABLE 12.1
Effect of Oxygen Pressure on Corrosion Rate of Iron in 3.5% Sodium Chloride Solution

Oxygen Pressure (atm)	Corrosion Rate (mm yr^{-1})
0.2	2.2
1	9.3
10	86.4
61	300

Source: Reprinted from *Corrosion Resistance of Metals and Alloys*, F. L. La Que and H. R. Capson, eds., Reinhold, 1963.

together constitute the corrosion process, proceed spontaneously or not? Such questions concerning the spontaneous occurrence of reactions fall within the scope of equilibrium thermodynamics.

Now, there are several ways in which thermodynamics can be used to answer the question at hand. For instance, one can make use of the relation between free-energy change and equilibrium potential to obtain the free-energy changes for the deelectronation and electronation reactions. The sum of the two free-energy changes yields the total free-energy change for the corrosion process

$$\Delta G = -nFV \tag{12.6}$$

If this total free-energy change is negative, then the corrosion of the metal will proceed spontaneously.

There is, however, a shortcut approach based on the potential vs. pH representation of equilibrium potentials. The approach is as follows: Suppose the $M^{n+} + ne \rightleftharpoons M$ reaction does not involve proton transfer. Its equilibrium potential is then independent of pH and can therefore be represented on the potential–pH diagram as a straight line parallel to the pH axis (Fig. 12.10). Next, one considers the electron acceptor A present in the solution in contact with the metal M and calculates the equilibrium potential for its reactions. Suppose it involves a proton transfer as well, i.e., $xA + mH^{+} + ne \rightleftharpoons yD + zH_2O$. Since this reaction involves both electron and proton
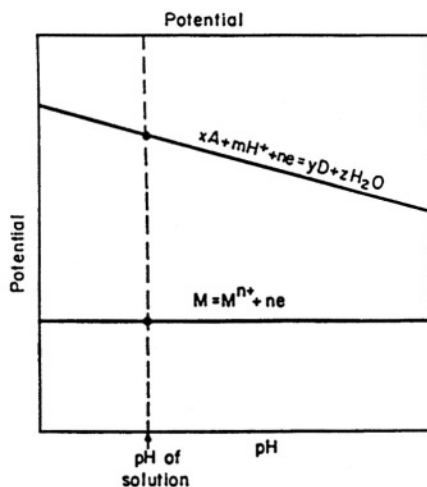

Fig. 12.10. The potential–pH diagram for the given system of the metal dissolution and electronation reaction reveals the tendency of the metal to corrode.

transfer, its equilibrium potential will vary with pH and can be represented as a straight line sloping downward in the potential–pH diagram.

Once one has a pH–potential diagram with lines drawn for the $M^{n+} + ne \rightleftharpoons M$ reaction and for the $xA + mH^+ + ne \rightleftharpoons yD + zH_2$ reaction, all one has to do is to draw a line perpendicular to the pH axis at the particular value of pH corresponding to that of the solution (Fig. 12.10). If that line intersects the $M^{n+} + ne \rightleftharpoons M$ line at a more negative value of potential than the $xA + mH^+ + ne \rightleftharpoons yD + zH_2O$ line, then a simple conclusion follows. The $M^{n+} + ne \rightleftharpoons M$ reaction will tend to run spontaneously in the deelectronation direction and produce M^{n+} from M (i.e., dissolution), and the other reaction will tend to proceed spontaneously as an electronation reaction (and thus absorb electrons supplied during the deelectronation of the metal) if a path is provided for the electron flow from the sink for the deelectronation reaction to the source for the electronation reaction. The metal M will be said to tend to corrode spontaneously.

On this basis, it is clear from Fig. 12.11 that if the solution in the mild-steel reaction vessel contains Fe^{2+} ions at a concentration of unit activity and if the pH = 2, the material of the vessel must tend to dissolve. It must, therefore, be rejected as an unsuitable material for holding a pH 2 solution.

Suppose, however, that the solution in the reaction vessel does not contain any ferrous ions. Then what is the concentration or activity that must be inserted in the Nernst expression for the equilibrium potential of the $Fe^{2+} + 2e \rightleftharpoons Fe$ reaction? What value should be used for $c_{Fe^{2+}}$ in
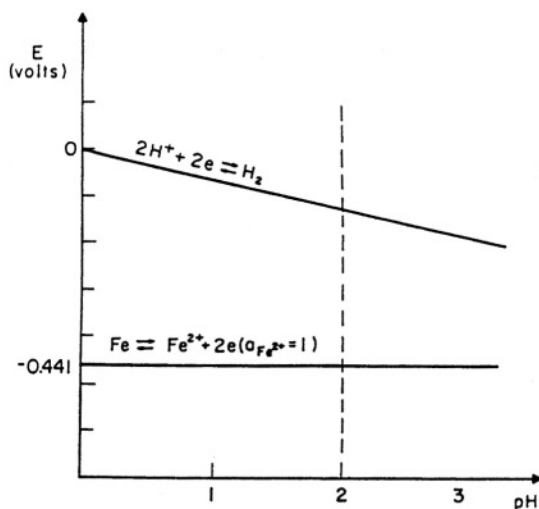

Fig. 12.11. The potential–pH diagram for iron immersed in a ferrous ion solution of unit activity.

$$E_{e,\text{Fe}^{2+}/\text{Fe}} = E^{\circ}_{\text{Fe}^{2+}/\text{Fe}} + \frac{RT}{2F} \ln c_{\text{Fe}^{2+}} \tag{12.7}$$

The obvious value to insert for $c_{\text{Fe}^{2+}}$ is zero in which case Nernst's equation for the electrode potential of the $\text{Fe}^{2+} + 2e \rightleftharpoons \text{Fe}$ equilibrium indicates a value that would be highly negative. Since this potential is negative with respect to the equilibrium potential of any possible electronation reaction, the iron will start dissolving and building Fe^{2+} concentration in the solution layer that is in contact with the electron-sink area of the iron. If this layer is stagnant and the Fe^{2+} ions are not removed by a chemical reaction, e.g., precipitation, the Fe^{2+} concentration will climb up from zero. Clearly, the Fe^{2+} ion concentration adjacent to the metal is determined by the amount of metal that has dissolved, and how fast this diffuses away.

How much iron must dissolve to attain a given Fe^{2+} concentration? A ferrous ion concentration of 10^{-6} mol liter^{-1} corresponds to the dissolution of about 0.06 mg of iron per liter of solution. Hence, a concentration of less than 10^{-6} mol liter^{-1} (e.g., 10^{-8} mol liter^{-1}) corresponds to the dissolution of about a microgram of iron per liter of solution in contact with the metal. On the other hand, a ferrous ion concentration of more than 10^{-6} mol liter^{-1} (e.g., 10^{-4} mol liter^{-1}), requires the dissolution of a few milligrams of iron per liter of solution in contact with it, i.e., the dissolution of a significant quantity of iron.

In view of these considerations, one can adopt practical and reasonable, though arbitrary, criterion: A ferrous ion concentration of 10^{-6} mol liter^{-1} and higher implies the occurrence of "considerable" dissolution, i.e., of corrosion. With these considerations as background, it is conventional in using a potential–pH diagram for deciding whether a metal can possibly corrode, to calculate the equilibrium potential for the $M^{n+} + ne \rightleftharpoons M$ reaction for a metal–ion concentration of 10^{-6} mol liter^{-1}.

Coming back to the question of whether a mild-steel vessel will corrode in a solution not initially containing Fe^{2+} ions, the answer can now be found by examining the position of the $\text{Fe}^{2+} + 2e \rightleftharpoons \text{Fe}$ line drawn for a concentration of 10^{-6} mol liter^{-1} of Fe^{2+}. This line will be $(RT/2F) \times \ln 10^{-6}$ V below that shown in Fig. 12.11. When this line then gets into a region more negative than the hydrogen line, iron will corrode.

The corrosion of an iron vessel has been treated here only to make the discussion less abstract. A similar approach can be used to inspect the potential–pH diagrams of other metals and decide whether they tend to corrode spontaneously in solutions of a given pH.

12.1.6. Potential–pH (or Pourbaix) Diagrams: Uses and Abuses

It must not be imagined that the *ultimate* product of the metal-dissolution reaction is always an ionic species, e.g., $M \rightarrow M^{n+} + ne$. Often it is a solid oxide or hydroxide.

From the free-energy considerations, one can calculate the reversible potentials for a metal that is in equilibrium with its simple hydrated ions or with its soluble product of hydrolysis or with its insoluble oxide. Under the given conditions, the

preferred state is that which gives the most negative values of potential, and any other existing state would spontaneously turn into that preferred one.

Potential–pH diagrams are useful in this respect, also. They indicate the potential and pH conditions under which a solid product is thermodynamically stable (Fig. 12.12). The regions of the potential–pH diagram in which oxide or hydroxide formation receives thermodynamic approval arise as follows:

Consider the case of iron, and assume for the sake of argument that the *immediate* product of iron dissolution is ferrous ions. Now, the solution in contact with iron can dissolve ferrous ions only up to the limit that is given by applying the law of mass action to the reaction:

$$Fe(OH)_2 + 2H^+ \rightleftharpoons Fe^{2+} + 2H_2O \tag{12.8}$$
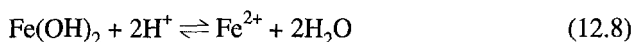
According to the law of mass action, for a constant concentration of $Fe(OH)_2$ in equilibrium with a solid phase,

$$\frac{c_{Fe^{2+}}}{c_{H^+}^2} = K = 10^{+13.29} \tag{12.9}$$
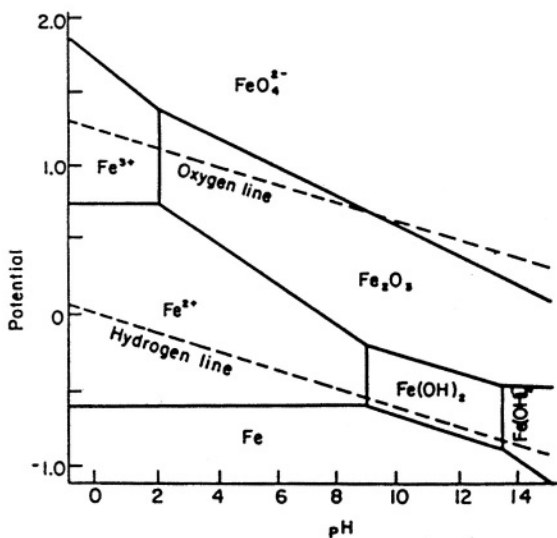
or

Fig. 12.12. Example of potential–pH diagram for a system with a solid phase as a dissolution product. Iron has a tendency to corrode at all pHs, but at pH $\geq$ 9 it forms $Fe(OH)_2$.

$$\log c_{Fe^{2+}} = 2 \log c_{H^+} + 13.29$$

$$= 13.29 - 2 \text{ pH} \tag{12.10}$$

If, as before, the minimum concentration of Fe^{2+} ions corresponding to a piece of corroding iron is arbitrarily taken as 10^{-6} mol liter^{-1}, then

$$\text{pH} = \frac{13.29 + 6}{2} = 9.6$$

which means that above a pH of 9.6, $Fe(OH)_2$ is stable. Since the $Fe(OH)_2/Fe^{2+}$ equilibrium depends only on pH and not on potential i.e., it is a pure proton-transfer reaction and does *not* involve electron transfer, the $Fe(OH)_2/Fe^{2+}$ equilibrium is shown on the potential–pH diagram as a vertical line parallel to the potential axis (see Fig. 12.12). Above the indicated pH value, the concentration of Fe^{2+} ions is governed by Eq. (12.10), and, hence, the potential changes with increasing pH with a slope of *RT/F*.

The question of how the corrosion of a metal is affected by the formation of a solid product of the dissolution reaction is a rather complex matter, which will be considered in due course. It is important, however, to stress one important point that is relevant not only to oxide formation. It often turns out that whereas the potential–pH diagram indicates that a particular hydroxide, e.g., can be formed only above a certain pH value, it is experimentally observed that the hydroxide is formed when the electrode is immersed in a solution with a much lower pH value. This apparent contradiction arises because (1) *the values of pH in a potential–pH diagram always refer to the solution in the immediate vicinity of the electrode* and (2) the local pH near an electrode can increase well above the bulk pH of the solution if the electronation reaction taking place at the electrode consumes hydrogen ions (e.g., $2H^+ + 2e \rightarrow H_2$) or generates hydroxyl ions (e.g., $O_2 + 2H_2O + 4e \rightarrow 4OH^-$).

In conclusion, therefore, potential–pH diagrams can be used to obtain yes or no answers on whether a particular corrosion process is thermodynamically possible. The diagrams provide a compact pictorial summary of the electron-transfer, proton-transfer, and electron-and-proton-transfer reactions that are favored on thermodynamic grounds when a metal is immersed in a particular solution. Yet, they should be used with caution. On the one hand, when a potential–pH diagram indicates that a particular metal is immune to corrosion, it is immune *provided* the pH in the close vicinity of the surface is what it is assumed to be. On the other hand, when the diagram indicates that a particular corrosion process can occur spontaneously, it does not mean that significant corrosion must actually be observed. For this to be so, the *rate* of corrosion must be appreciable, and one must at this stage refrain from making any predictions regarding the rate of corrosion since this cannot be done from a knowledge of the thermodynamics of the system alone. If one does not observe this caution, one can be led into serious errors. A classical example of the point under discussion is the case of lead in contact with aerated water. The potential–pH diagram (Fig. 12.13)
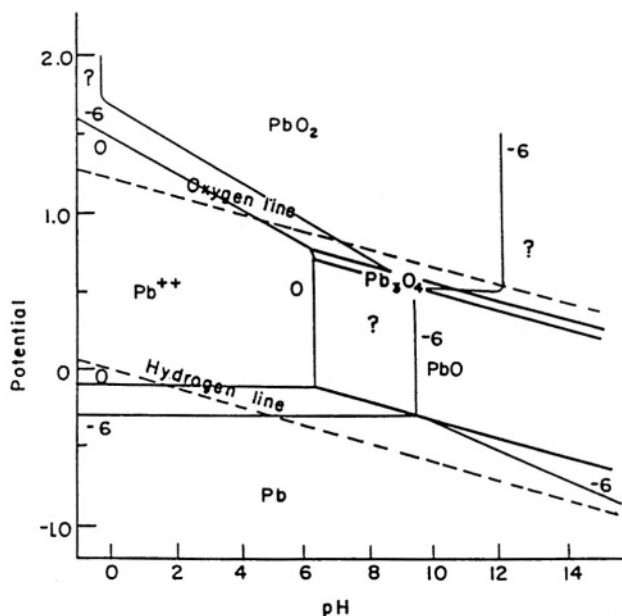
Fig. 12.13. The potential–pH diagram for lead.

shows the equilibrium potential for the $Pb^{2+} + 2e \rightleftharpoons Pb$ reaction for a Pb^{2+} concentration of $\sim 10^{-6}$ g liter^{-1} to be negative with respect to the equilibrium potential for the hydrogen-reduction reaction at a pH less than about 5. This implies a tendency on the part of lead to corrode in an aerated aqueous environment. In actual fact, however, the rate of corrosion is so negligible that lead is often used in pipes for carrying water.

Thermodynamics, therefore, defines a necessary, vital precondition for corrosion; it determines the direction in which an overall corrosion reaction will *tend*. But the determination of the rate and control of a corroding system can emerge only by a study of the *electrodics of corrosion*.

12.1.7. The Corrosion Current and the Corrosion Potential

Consider a system consisting of a metal corroding in an electrolyte. The corrosion process involves a metal-dissolution deelectronation (anodic) reaction at electron-sink areas on the metal and an electronation (cathodic) reaction at electron-source areas. (This picture is applicable to a metal's corroding by a Wagner–Traud mechanism provided one imagines the sink and source areas shrunk to atomic-sized dimensions and considers the situation at one instant of time.)

A corroding metal, it has been pointed out, is equivalent to a short-circuited energy-producing cell with the following specifications: The electron sink (anodic) and electron source (cathodic) areas of the equivalent energy-producing cell are chosen

equal to the corresponding areas on the corroding metal. Thus, the total metal-dissolution *current* I_M and electronation *current* I_{SO} (not current densities) on the corroding metal are equal in magnitude but opposite in sign, just as they are in an energy-producing cell,

$$I_M = -I_{SO} \tag{12.11}$$

The rate of corrosion of the metal is obviously given directly by the rate of metal dissolution: hence, the corrosion current I_{corr} is equal to the metal-dissolution current I_M:

$$I_{corr} = I_M = -I_{SO} \tag{12.12}$$

Apart from this important feature of a corroding system, there is another characteristic that arises from the *short-circuit* condition of the corrosion cell and of the equivalent cell. (It will be recalled that the electron sources and sinks in the corroding metal are *internally* short-circuited; the two electrodes in the equivalent cell are *externally* short-circuited.) The total potential difference V across the equivalent cell is zero. But this cell potential is composed of the absolute potential differences across the interfaces at the two electrodes and the potential drop IR in the electrolyte:

$$V = 0 = \Delta\phi_{SO} - \Delta\phi_M + IR \tag{12.13}$$

where $\Delta\phi_{SO}$ is the metal–solution potential difference at the electron source electrode (cathode) and $\Delta\phi_M$ is the corresponding quantity at the electron sink electrode (anode).

Now assume that $IR \approx 0$. The assumption requires that the interelectrode distance be negligibly small, that the electrolyte be sufficiently conducting, and that there be no high-resistance oxide films on the electrodes. Under these circumstances,

$$\Delta\phi_{SO} - \Delta\phi_M \approx 0 \quad \text{or} \quad \Delta\phi_{SO} \approx \Delta\phi_M \tag{12.14}$$

Thus, when $IR \approx 0$, the potential difference across the metal/electrolyte interface at the electron source electrode of the short-circuited equivalent cell is virtually equal to that at the electron sink electrode.

What is the validity of the assumption $IR \approx 0$ in the case of a corroding metal? If the metal is homogeneous and is corroding by a Wagner–Traud mechanism, the sink and source areas are separated at any one instant by a distance on the order of a few angstroms. Further, the sink and source areas are shifting around with time and smearing out the negligible potential differences in the solution adjacent to these areas. Thus, $IR = 0$ is almost exactly true.

If the metal has heterogeneities and is corroding by local-cell action, the validity of $IR \approx 0$ depends upon the separation of the sink and source areas and upon the conductivity of the electrolyte. Now, there are special circumstances in which the distance apart of the sink and source areas is considerable (on the order of centimeters).

For these situations, $IR \approx 0$ and $\Delta\phi_{SO} \approx \Delta\phi_M$, which implies a difference in the metal–electrolyte potential difference at electron-source and -sink areas. In general, however, the sink-to-source distance is on the order of microns or less, in which case the conducting path in the solution and therefore IR becomes negligible. Thus, the $\Delta\phi_{SO}$ is virtually equal to $\Delta\phi_M$, and any negligible difference that exists occurs over distances too small to be resolved by the probe used to measure the potential difference between the metal and the solution (Fig. 12.14).

This uniform potential difference across the interface between a corroding metal and its electrolytic environment may be termed the *corrosion potential* $\Delta\phi_{corr}$; it is considered to be given by

$$\Delta\phi_{corr} = \Delta\phi_{SO} = \Delta\phi_M \qquad (12.15)$$

It follows that the corrosion potential on a heterogeneous metal corroding by local-cell action is *virtually equal* to the mixed potential at an electrode on which electronation and deelectronation reactions are occurring on spatially separated sinks and sources and is *identical* to a mixed potential when the metal is corroding homogeneously by a Wagner–Traud mechanism. The concept of the corrosion current I_{corr} and the corrosion potential $\Delta\phi_{corr}$ will now be treated quantitatively.
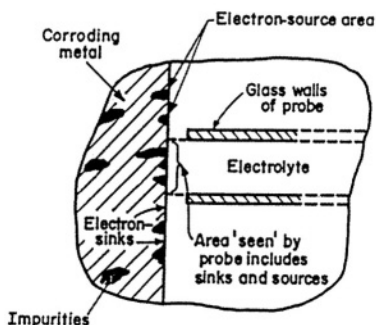

Fig. 12.14. In heterogeneous systems, the distances between source (e.g. metallic inclusions) and sink points are so small that a probe usually cannot detect any potential difference in the electrolyte between them.

12.1.8. The Basic Electrodics of Corrosion in the Absence of Oxide Films

Two fundamental ideas have been developed. First, the rate of corrosion, a quantity of great practical significance, is given by the corrosion current I_{corr}, which is equal to the metal-dissolution, deelectronation current I_M and to the negative of the electronation (cathodic) current I_{SO} at the electron-source areas, i.e.,

$$I_{corr} = I_M = -I_{SO} \qquad (12.12)$$

Since the metal-dissolution current is equal to the product of the corresponding current density i_M times the sink area A_M, one can write:

$$I_{corr} = I_M = A_M i_M \qquad (12.16)$$

and, similarly,

$$I_{corr} = -I_{SO} = -A_{SO} i_{SO} \qquad (12.17)$$

Second, there is a uniform potential difference, namely, the corrosion potential $\Delta\phi_{corr}$, all over the surface of the corroding metal. It is this corrosion potential that is associated with both the metal-dissolution and electronation currents, i.e.,

$$\Delta\phi_{corr} = \Delta\phi_M = \Delta\phi_{SO} \qquad (12.15)$$

To obtain quantitative expressions for the corrosion current and the corrosion potential, one has to substitute the proper expression for the metal-dissolution- and electronation-current densities. If no oxide films form on the surface of the corroding metal and neither of the current densities is controlled by mass transport, i.e., there is no concentration overpotential, one can insert the Butler–Volmer expression for the deelectronation- and electronation-current densities. Thus,

$$I_{corr} = I_M = A_M i_M \qquad (12.16)$$

$$= A_M i_{0,M} \left[\exp\left(\frac{\overleftarrow{a}_M F}{RT}\, \eta_M \right) - \exp\left(-\frac{\overrightarrow{\alpha}_M F}{RT}\, \eta_M \right) \right] \qquad (12.18)$$

Now, the overpotential η_M is equal to the potential difference at the electron sink areas, i.e., the corrosion potential $\Delta\phi_{corr}$ minus the equilibrium potential $\Delta\phi_{eM}$ for the metal-dissolution reaction $M^{n+} + ne \rightleftharpoons M$, i.e.,

$$\eta_M = \Delta\phi_{corr} - \Delta\phi_{eM} \qquad (12.19)$$

Further, since $i_{0,\text{M}}$ is the exchange-current density for the $\text{M}^{n+} + ne = \text{M}$ reaction and A_M is the area over which this reaction occurs, the product of $i_{0,\text{M}}$ and A_M must be the exchange *current*, i.e.,

$$I_{0,\text{M}} = A_\text{M} i_{0,\text{M}} \tag{12.20}$$

Finally, as was done in the treatment of electrochemical cells (see Section 7.13.5), one can use the following notation:

$$\overleftarrow{\lambda}_\text{M} = \frac{RT}{\overleftarrow{\alpha}_\text{M} F} \quad \text{and} \quad \overrightarrow{\lambda}_\text{M} = \frac{RT}{\overrightarrow{\alpha}_\text{M} F} \tag{12.21}$$

where $\overleftarrow{\lambda}_\text{M}$ and $\overrightarrow{\lambda}_\text{M}$ are the $\eta_\text{M} - \log i_\text{M}$ Tafel slopes for the deelectronation and electronation directions of the $\text{M}^{n+} + ne \rightleftharpoons \text{M}$ reaction.

In view of Eqs. (12.19) to (12.21), the expression (12.18) for the corrosion current becomes

$$I_\text{corr} = I_{0,\text{M}} \left[\exp \frac{\Delta\phi_\text{corr} - \Delta\phi_{e,\text{M}}}{\overleftarrow{\lambda}_\text{M}} - \exp\left(- \Delta\phi_\text{corr} - \frac{\Delta\phi_{e,\text{M}}}{\overrightarrow{\lambda}_\text{M}}\right) \right] \tag{12.22}$$

Similarly, one can write for the relation between the corrosion current and the electronation current at the electron-source area:

$$I_\text{corr} = -I_\text{SO}$$

$$= -A_\text{SO} i_{0,\text{SO}} \left[\exp \frac{\overleftarrow{\alpha}_\text{SO} F \eta_\text{SO}}{RT} - \exp\left(-\frac{\overrightarrow{\alpha}_\text{SO} F \eta_\text{SO}}{RT}\right) \right]$$

$$= I_{0,\text{SO}} \left[\exp\left(-\frac{\Delta\phi_\text{corr} - \Delta\phi_{e,\text{SO}}}{\overrightarrow{\lambda}_\text{SO}}\right) - \exp \frac{\Delta\phi_\text{corr} - \Delta\phi_{e,\text{SO}}}{\overleftarrow{\lambda}_\text{SO}} \right] \tag{12.23}$$

From Eqs. (12.22) and (12.23), it is clear that the corrosion current depends upon the exchange currents (i.e., available areas and exchange-current densities), Tafel slopes, and equilibrium potentials for both the metal-dissolution and electronation reactions. To obtain an explicit expression for the corrosion current [cf. Eq. (12.22)], one has first to solve Eqs. (12.22) and (12.23) for $\Delta\phi_\text{corr}$. If, however, simplifying assumptions are not made, the algebra becomes unwieldy and leads to highly cumbersome equations.

One such simplifying assumption is

$$\overleftarrow{\alpha}_\text{M} = \overrightarrow{\alpha}_\text{M} = \overleftarrow{\alpha}_\text{SO} = \overrightarrow{\alpha}_\text{SO} = \frac{1}{2} \tag{12.24}$$

i.e.,

$$\overleftarrow{\lambda}_M = \overrightarrow{\lambda}_M = \overleftarrow{\lambda}_{SO} = \overrightarrow{\lambda}_{SO} = \frac{2RT}{F} \tag{12.25}$$

If, making use of this assumption, one divides the two expressions (12.22) and (12.23) for I_{corr} by exp $(- F\Delta\phi_{corr}/2RT)$ and equates them, the result is

$$I_{0,M}(e^{F\Delta\phi_{corr}/RT}e^{-F\Delta\phi_{e,M}/2RT} - e^{F\Delta\phi_{e,M}/2RT})$$

$$= I_{0,SO}(e^{F\Delta\phi_{e,SO}/2RT} - e^{F\Delta\phi_{corr}/RT}e^{-F\Delta\phi_{e,SO}/2RT}) \tag{12.26}$$

An evaluation of $\Delta\phi_{corr}$ from this expression shows that

$$\Delta\phi_{corr} = \frac{RT}{F} \ln\left[\frac{I_{0,SO} \exp(F\Delta\phi_{e,SO}/2RT) + I_{0,M} \exp(F\Delta\phi_{e,M}/2RT)}{I_{0,SO} \exp(- F\Delta\phi_{e,SO}/2RT) + I_{0,M} \exp(- F\Delta\phi_{e,M}/2RT)} \right] \tag{12.27}$$

Although obtained under the simplifying assumption of $\overleftarrow{a}_M = \overrightarrow{a}_M = \overleftarrow{a}_{SO} = \overrightarrow{a}_{SO} = \frac{1}{2}$, this equation brings out a simple characteristic of the corrosion potential; it approximates being near the equilibrium potential for the metal-dissolution reaction or near the equilibrium potential of the electronation reaction, depending upon whether the exchange current at the sink areas is much greater than the exchange current at the source areas or vice versa. In symbols, if $I_{0,M} >> I_{0,SO}$, then $\Delta\phi_{corr} \approx \Delta\phi_{e,M}$, and, if $I_{0,SO} >> I_{0,M}$, then $\Delta\phi_{corr} \approx \Delta\phi_{e,SO}$.

The expression (12.27) for the corrosion potential can be introduced into Eq. (12.22) and thus, an explicit result for the corrosion current can be obtained. But the resulting equation is quite cumbersome and therefore a simpler equation will be derived by assuming that overpotentials are sufficiently large that the high-field approximation of the Butler–Volmer equation can be used for the electronation- and deelectronation-current densities. Thus, Eqs. (12.22) and (12.23) become

$$I_{corr} = I_{0,M} \exp\frac{\Delta\phi_{corr} - \Delta\phi_{e,M}}{\overleftarrow{\lambda}_M} = I_{0,SO} \exp\left(- \frac{\Delta\phi_{corr} - \Delta\phi_{e,SO}}{\overrightarrow{\lambda}_{SO}}\right) \tag{12.28}$$

Hence,

$$\exp\frac{\Delta\phi_{corr}}{\overleftarrow{\lambda}_M} = \left(\frac{I_{0,SO}}{I_{0,M}}\right)^{\overrightarrow{\lambda}_{SO}/(\overleftarrow{\lambda}_M + \overrightarrow{\lambda}_{SO})} \exp\left[\overleftarrow{\lambda}_M\Delta\phi_{e,SO} + \frac{\overrightarrow{\lambda}_{SO}\Delta\phi_{e,M}}{\overleftarrow{\lambda}_M(\overleftarrow{\lambda}_M + \overrightarrow{\lambda}_{SO})}\right] \tag{12.29}$$

and, therefore,

$$I_{corr} = I_{0,M}^{\overleftarrow{\lambda}_M/(\overleftarrow{\lambda}_M + \overrightarrow{\lambda}_{SO})} I_{0,SO}^{\overrightarrow{\lambda}_{SO}/(\overleftarrow{\lambda}_M + \overrightarrow{\lambda}_{SO})} \exp\left[\frac{\Delta\phi_{e,SO} - \Delta\phi_{e,M}}{\overleftarrow{\lambda}_M + \overrightarrow{\lambda}_{SO}}\right] \tag{12.30}$$

The dependence of the corrosion current (the rate at which a metal destroys itself) on the exchange currents, Tafel slopes, and equilibrium potentials of the metal-dissolution and electronation reactions is clearly brought out in this expression. In general, the more positive the equilibrium potential of the electronation reaction is with respect to the equilibrium potential for the $M^{n+} + ne \rightleftharpoons M$ reaction and the larger the exchange currents (areas times exchange-current densities) are, the greater is the rate of corrosion. The Tafel slopes also enter the picture; high slopes diminish the enhancing effect that the exponential term has on the rate of the corrosive attack on the metal.

A simpler form of Eq. (12.30) can be determined by setting

$$\overleftarrow{\lambda}_M = \overleftarrow{\lambda}_{SO} = \frac{2RT}{F} \tag{12.25}$$

in which case one gets

$$I_{corr} = (I_{0,M}I_{0,SO})^{\frac{1}{2}} \exp\frac{F(\Delta\phi_{e,SO} - \Delta\phi_{e,M})}{4RT} \tag{12.31}$$

and if the potentials are written as relative potentials on the standard hydrogen scale, Eq. (12.31) becomes

$$I_{corr} = (I_{0,M}I_{0,SO})^{\frac{1}{2}} \exp\frac{F(E_{e,SO} - E_{e,M})}{4RT} \tag{12.32}$$

This approximate and special-case equation brings out the role of the exchange currents and the equilibrium potentials in determining the corrosion rate.

What has been presented above is a very elementary account of corrosion under super-ideal conditions. In a few cases, it does give a fairly good agreement with the observed rates of corrosion. Yet, in real systems, corrosion is nearly always too complex a phenomenon for the above simple treatment to be directly applicable. The simple version would be valid *if* there were no oxide films, *if* there were a negligible IR drop in the solution, *if* the corrosion potential $\Delta\phi_{corr}$ settled down to a value such that the high-field approximations [cf. Eq. (12.28)] could be applied, and *if* the transfer coefficients of the metal-dissolution and electronation reactions were $\frac{1}{2}$ [cf. Eq. (12.25)]. However, the point of an introductory treatment is not to treat the details and the complex realities, but to present the idealized essence about an electrochemical mechanism that has substantial effects in the everyday world.

12.1.9. An Understanding of Corrosion in Terms of Evans Diagrams

Most of the factors affecting the rate of corrosion can be understood from a graphical superposition of the current-potential curves for the metal-dissolution and electronation reactions. The principle of the graphical superposition method is straightforward.

Consider the metal-dissolution reaction $M^{n+} + ne \rightleftharpoons M$. One can construct a curve (Fig. (12.15) for the variation of the potential of an M electrode with the deelectronation current crossing the electrode/electrolyte interface. This curve can be obtained either experimentally or from a knowledge of the parameters that determine the overpotential associated with the deelectronation-current density. For concentration overpotential, this parameter is the limiting current density, and, for activation over-potential, the parameters are the exchange-current density and the transfer coefficients. On the same diagram, one can then superpose a curve (Fig. 12.16) for the variation of the potential of the M electrode with the current associated with the electronation of electron acceptors present in the electrolyte. The current at which the metal dissolution and electronation are equal is, in fact, the corrosion current (cf. Fig. 12.16). The potential corresponding to the corrosion current is the corrosion potential. If one uses only the *magnitude* of the deelectronation and electronation currents in the construction of the $\Delta\phi$ versus I curves, one has what is known as an *Evans* type of diagram (Fig. 12.17).

The particular form of the Evans diagram obtained depends upon the current-potential curves for the metal-dissolution and electronation reactions. Some of the common diagrams are shown in Figs. 12.18 to 12.20. They cover situations in which
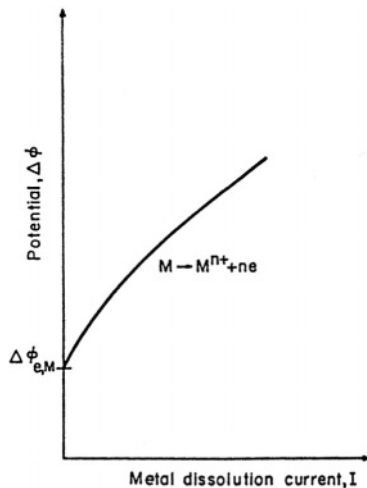

Fig. 12.15. The potential–current relation for the metal dissolution re-action.
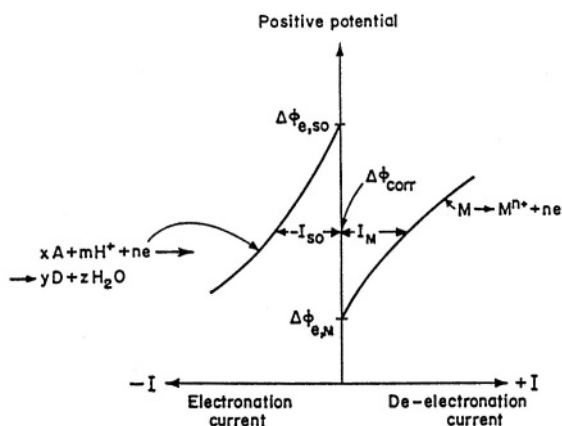
Fig. 12.16. The potential–current relation for the two reactions occurring at a corroding interface. The corrosion current and corrosion potential are defined by the point on the diagram at which the two currents I_M and I_{SO} are equal.

the exchange current for the metal-dissolution reaction is much greater than that for the electronation, i.e., $I_{0,M} >> I_{0,SO}$ [Fig. 12.18(a)] or in which $I_{0,SO} >> I_{0,M}$ [Fig. 12.18(b)]. Evans diagrams can also be used to bring out the influence of Tafel slopes [Fig. 12.19(a)], the influence of equilibrium potentials [Fig. 12.19(b)] and the effect
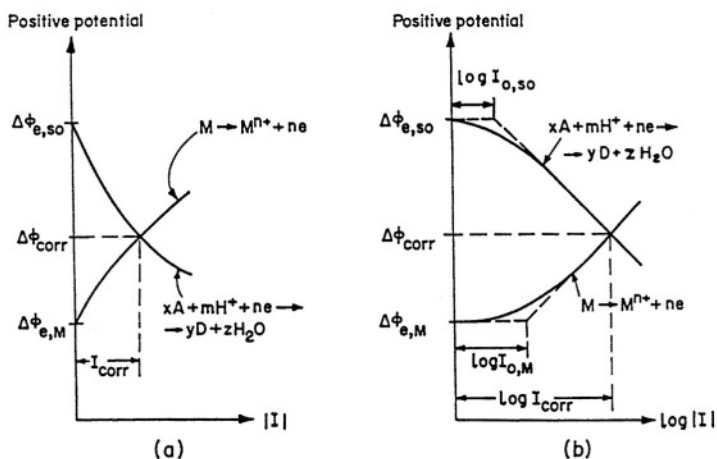

Fig. 12.17. The Evans diagrams are plots of the potentials of the two reactions (a) vs. the magnitude of the two currents or (b) vs. their logarithms. The intersections of the curves define the corrosion current and corrosion potential.
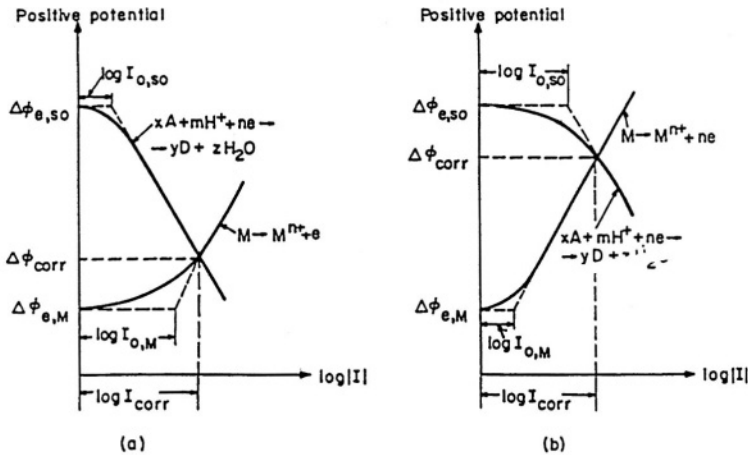
Fig. 12.18. Changing the exchange current densities produces a shift of the corrosion potential from a medium value toward the equilibrium potential of (a) the metal dissolution or (b) the electronation reaction.

of mass-transport control on the electronation current [Fig. 12.19(c)]. The effect of an IR drop in the electrolyte between the electron-sink (anodic) and electron-source (cathodic) areas can also be represented in an Evans diagram (Fig. 12.20), which then shows the inequality of the metal-solution potential difference at the two areas, i.e., the anodic (or deelectronation) and the cathodic (or electronation) areas.

12.1.10. How Corrosion Rates Are Measured

There are basically three kinds of approaches to the measurement of corrosion. Although the mechanism of corrosion is certainly electrochemical, only one of the methods available uses an ostensibly electrochemical technique.

12.1.10.1. *Method 1: The Weight-Loss Method.*   Here, samples of the metal or alloy, the corrosion rate of which is to be determined, are suspended in a solution (e.g., sea water) containing the appropriate aggressive ions (e.g., Cl^-) that may be met in practice. The weight of the sample is measured at regular intervals over a long period, e.g., a year. Assuming that the change in weight represents only a loss of metal to the solution, corrosion in the simplest sense, it can be converted to $mol\ cm^{-2}\ s^{-1}$ (= rate) or a corrosion current in amperes cm^{-2} from i_{corr}/nF = rate in $mol\ cm^{-2}$. Corrosion engineers may re-express the measurement in microinches $year^{-1}$.

This weight loss method is seen by engineers as "real" (representing the overall result of corrosion, however it occurs). The long times involved represent reality more than do short time measurements (see Method 2), which might be expected over many years in a practical situation.
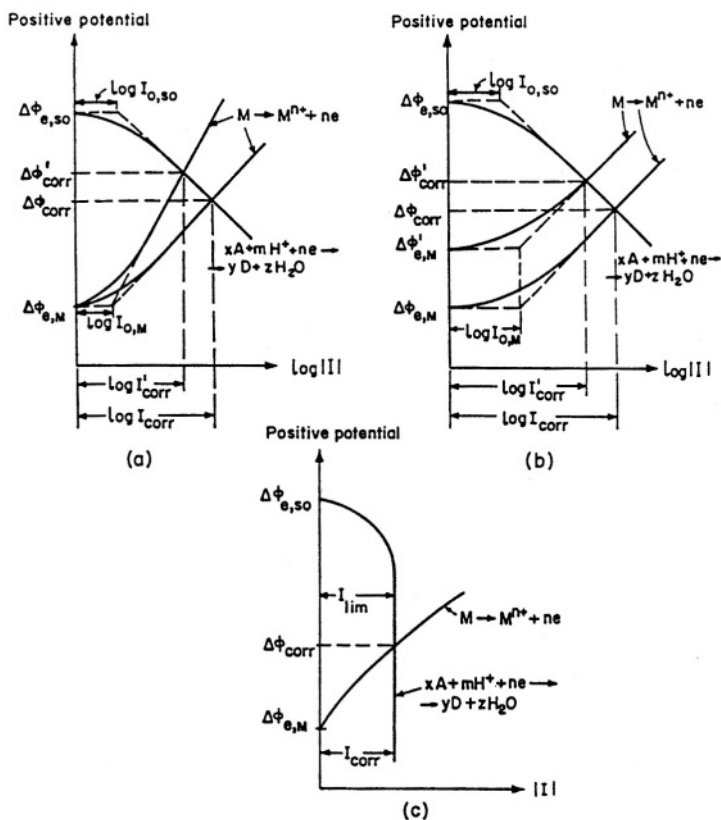
Fig. 12.19. The effect of (a) the Tafel slope, (b) the equilibrium potential of the metal dissolution, or (c) the transport difficulties of the electron acceptor are reflected in the Evans diagram.

On the other hand, such a long-term approach won't do if one wishes to determine relatively quickly the corrosivities of a number of metals in a certain ambient; or one metal in the presence of a series of corrosion inhibitors. Another limitation arises from the fact that the most dangerous corrosion (that which sometimes causes the collapse of bridges or buildings) involves internal cracking of the metal (Section 12.6.5) so that, with little loss to the solution, a dangerous loss of strength in the metal can occur.

12.1.10.2. *Method 2: Electrochemical Approach.* The equations (12.23) developed above for the anodic and cathodic reactions in corrosion, which, when equal, lead to a steady-state corrosion current, contain two quantities that are better simplified in order to present the basis of the electrochemical method for measuring corrosion. Thus, the Galvani potential differences used there can be converted to
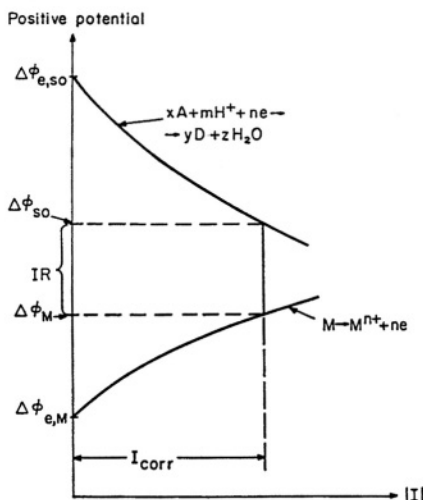
Fig. 12.20. If there is some potential drop in the solution (IR), the two potential–current curves do not intersect at the corrosion current.

electrode potentials by using the relation $V = \Delta\phi +$ const. Further, instead of the accurate equation (12.22), which involves both forward and backward rates of the anodic reaction, only the forward reaction will be counted here, an assumption which implies that the corrosion potential is at least RT/F anodic to its reversible potential. A similar assumption is made for the cathodic current. Then:

$$i_{anodic} = i_{corr}(e^{+\alpha_a(V-V_{corr})F/kT} - e^{-\alpha_a(V-V_{corr})F/RT}) \qquad (12.33)$$

Such an equation represents the anodic current density that would be measured if a metal with a steady-state corrosion[2] current of i_{corr} A cm^{-2} is polarized to a potential V, positive to the corrosion potential, V_{corr}.

---

[2]It is worth noting here that we are dealing with a steady-state situation, not with an equilibrium one. Thus, although Eqs. (12.33) and (12.34) bear a formal resemblance to the Butler–Volmer equation (7.24), there is an important difference. In the latter, the i_0 situation arises at zero overpotential when the electrode potential, V, is equal to the thermodynamically reversible potential for some equilibrium such as $M^+ + e_0^- \rightleftharpoons M$. However, in the analogous equation (12.34), we are dealing with the partial currents, one anodic (that of the true metal dissolution, the very essence of corrosion), and the other cathodic, the "depolarizing reaction" that mops up the electrons injected back into the metal when the cation is formed from the metal and thus allows corrosion to continue at a steady rate. Neither of these reactions occurs at equilibrium, but at V_{corr}, the corrosion potential. Hence, Eqs. (12.33) and (12.34) cannot represent an equilibrium, and the i_{corr} that we want to determine by using them is only *analogous* to the equilibrium exchange current density of the Butler–Volmer equation. In fact, it represents a *steady-state* situation made up of two equal electron streams, each opposite to the other in direction.

Likewise,

$$i_{\text{cathodic}} = i_{\text{corr}}(e^{-\alpha_c(V-V_{\text{corr}})F/RT} - e^{\alpha_a(V-V_{\text{corr}})F/RT}) \tag{12.34}$$

represents the cathodic current density that would be expected if the potential of the corroding metal were shifted away from the corrosion potential in the cathodic direction.

These equations can be understood by comparison with the Butler–Volmer equation for a single reaction [cf. Eq. (7.23)]. The overpotential, η is $V - V_{\text{rev}}$, where V_{rev} is the thermodynamic reversible potential of the one reaction; and i_0 is the exchange current density of that reaction. Similarly, for a corroding metal, the anodic and cathodic branches behave in a way analogous to those of the metal at which a single reaction is occurring. However, the cathodic reaction for the anodically dissolving metal reaction is neglected and replaced by cathodic current for the partner reaction in which the anodic current is neglected.

Figure 12.21 shows these two equations (the result of anodic and cathodic biasing, respectively, away from the corrosion potential). Now, if the exponents in Eq. (12.23) are sufficiently small, so that the corresponding exponentials can be expanded linearly, one obtains, e.g., for the anodic current density:

$$i = i_{\text{corr}} \frac{F}{RT}(\alpha_{\text{cathodic}} + \alpha_{\text{anodic}})(V - V_{\text{corr}}) \tag{12.35}$$

or

$$i_{\text{corr}} = \frac{RT}{F(\alpha_{\text{cathodic}} + \alpha_{\text{anodic}})}\left(\frac{\partial i}{\partial V}\right) \tag{12.36}$$

The range of potentials in which the linear relation (12.35) is valid is given by

$$(\alpha_{\text{cathodic}} + \alpha_{\text{anodic}}) \frac{F}{RT}(V - V_{\text{corr}}) < 1 \tag{12.37}$$

The terms α_{anodic} and α_{cathodic} are the transfer coefficients [Eq. (7.143)] for the anodic and cathodic components of the corrosion reaction, respectively. Their values will depend upon the reactions making up the corrosion situation. If one assumes that a hydrogen evolution rate controlled by charge transfer is the cathodic reaction, $\alpha_{\text{cathodic}} = 1/2$; and if, e.g., the metal dissolution is controlled by charge transfer to form a divalent cation, $\alpha_{\text{anodic}} = 2$. Then, from (12.37), the maximum value of $V - V_{\text{corr}}$ allowable for the approximation of Eq. (12.35) is

$$(V - V_{\text{corr}}) < \frac{RT}{(\alpha_{\text{cathodic}} + \alpha_{\text{anodic}})F} = \frac{0.026}{2\,1/2} = 0.01$$

for 25 °C.

Thus, to determine the corrosion current, i_{corr}, one displaces the potential of the corroding metal in a small range only up to 10 mV, and obtains $(\partial i/\partial V)$, the slope of the linear i–V relation near the corrosion potential.

The electrochemical approach to the measurement of the corrosion rate, i_{corr} (in A cm^{-2}) was originated by Stern and Geary (1957). Originally (see Fig. 12.21), the biasing of the potential was carried out to values outside the linear region derived above, i.e., into the Tafel region in which $\log i_{anodic}$ and $\log i_{cathodic}$ (biasing on the anodic and cathodic sides respectively) are linear (Fig. 12.21). Then, extrapolating them linearly to V_{corr}, the corresponding $\log i$ value there is $\log i_{corr}$.

However, the linear method (i.e., a small displacement of potential) is better. Thus, going out far from V_{corr} on the anodic side may run into a region in which the metal forms oxide films; and on the cathodic side, the evolution of H_2 could interfere with the anodic dissolution current, which could confusingly lead to an erroneous contribution (via $H_2 \rightarrow 2H^+ + 2e$) to the anodic dissolution reaction.
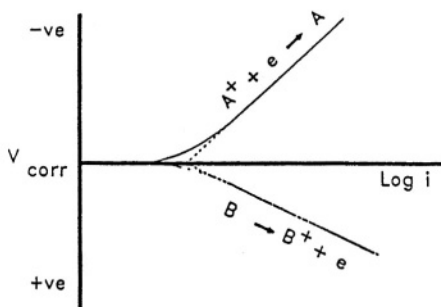

Fig. 12.21. The cathodic and anodic Tafel lines at two different electrode reactions. At V_{corr} there is a steady state. No net current passes, but there is a steady production of A (e.g., H_2) and B^+ (the metal, corroding) in a net electrochemical reaction that appears to be chemical (no net current flows to an outside circuit). Note the difference in these diagrams from the Evans–Hoar diagrams, which show corrosion is spontaneous and drives itself. The difference is similar to that between an electrochemical reaction and a fuel cell. (Reprinted from J. O'M. Bockris and S. N. M. Kahn, *Surface Electro Chemistry*, Fig. 8.1, p. 747 Plenum, 1993).

12.1.11. Impedance Bridge Version of the Stern–Geary Approach

It was shown earlier (Section 7.2.4) that for a simple one-reaction situation, the resistance of the interface is given by RT/i_0F, where i_0 is the exchange current $\eta < RT/F$. It follows that if the equilibrium of, e.g., $M^{2+} + ze_0^- \rightarrow M$ and $M \rightarrow M^{2+} + ze$ is replaced, e.g., by $A^+ + e \rightarrow A$ (where A is H_3O^+ or O_2) and $M \rightarrow M^+ + e$, when these two reactions are equal in rate, then

$$R_{corr} = \frac{RT}{i_{corr}F} \tag{12.38}$$

where i_{corr} is the steady-state corrosion rate. Thus, if a corroding metal is placed in an impedance bulge, and the resistance of the interface, R_{corr}, is determined, one can determine the rate of corrosion, i_{corr}, from Eq. (12.38).

The electrochemical approach has the advantage of speed and relative simplicity. The disadvantage is that one obtains the corrosion rate under the conditions chosen—a fresh electrode and solution—i.e., corrosion in the short term. Real corrosion situations are more complex. At longer times, the metal becomes partly covered with an oxide and other coatings; the solution or moisture film contains components not there in a laboratory situation. However, the Stern–Geary electrochemical approach allows at least a relative determination of the corrosion rate for a series of situations. It is simple and it is fast.

12.1.12. Other Methods

In spite of the fact that corrosion measurements have been made for a century and electrochemical methods used for about a half century, there is still a need for the development of new methods which would, e.g., predict long-term corrosion rates and that of internal corrosion that could lead to disastrous breakdowns.

Further, the above methods measure the equivalent global corrosion rate of a large area. Corrosion in reality is often local, i.e., takes place only at tiny areas of the surface. It is necessary to have a local probe, not only of the rate, but of the mechanical properties. Something about this approach will be given in the discussion of local corrosion (Section 12.5).

12.1.13. The Mechanisms of the Corrosion Reactions Involving the Dissolution of Iron

Because iron and alloys from iron (i.e., steel) are the principal metals used in construction, corrosion reactions involving them are of greater practical importance than the corresponding reactions on other metals (although those on aluminum are increasingly important).
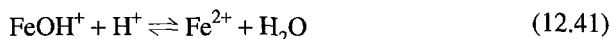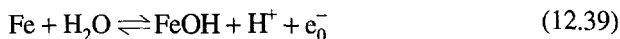
The discussion here concerns two subsections. In the first, the mechanism of the dissolution of Fe is discussed and in the second, those of the two most usual partner

cathodic reactions, hydrogen evolution for an acid solution and O_2 reduction for an alkaline solution.

Because mechanistic work on corrosion often has to depend on knowledge of the mechanism of the dissolution of Fe, it would be encouraging to be able to think that this important mechanism was clear and simple. The truth is that work during the past 30 years has shown the mechanism to be rather labile. Thus, unstressed iron dissolves anodically according to one sequence of reactions but after it is stressed (all other conditions being the same), the rate-determining step (rds) changes. This complicates the question of the relevant rds, which may evidently be different on various parts of a surface, depending on the local stress.

12.1.14. Something about the Mechanism of the Anodic Dissolution of Iron

For the restricted conditions of unstressed iron in $0 < pH < 6$, there are two main diagnostic results that suggest what is the most well-known mechanism for iron dissolution. Thus, under the conditions stated, the Tafel constant, b_{anodic}, is found to be $2RT/3F$ and the cathodic[3] slope, $b_{cathodic}$ is $-2RT/F$. Surprisingly, the reaction orders with respect to α_{OH^-} are 1 for both the anodic ($Fe \rightarrow Fe^{2+} + 2e$) and the reverse cathodic reactions. For the latter, the actual experimental reaction order found was 0.8, but it is usually taken as 1. A mechanism that fits these facts is

$$Fe + H_2O \rightleftharpoons FeOH + H^+ + e_0^- \tag{12.39}$$

$$FeOH \xrightarrow{rds} FeOH^+ + e_0^- \tag{12.40}$$

$$FeOH^+ + H^+ \rightleftharpoons Fe^{2+} + H_2O \tag{12.41}$$

Qualitatively, it is clear that such a mechanism might explain the unexpected pH dependence of the reaction rate. Then, with $^M\Delta^S = V + const$,

$$\theta_{FeOH} = \frac{K_1}{c_H^+} e^{VF/RT} = \frac{K_1}{K_w} C_{OH^-} e^{VF/RT}$$

with the rds as (12.40):

$$i_{anodic,Fe} = 2Fk_2Ke_{OH^-}e^{(1+B)VF/RT} \tag{12.42}$$

---

[3]The deposition of Fe from acid solutions is complicated by the co-deposition of hydrogen, and this latter co-reaction was measured in the special cell shown in Fig. 12.22. The co-deposited hydrogen was evolved in a solution presaturated with H_2. The sensitivity was greater than 10^{-3} ml H_2.
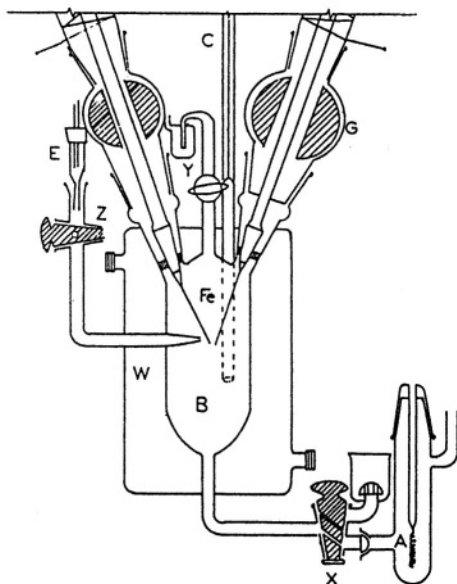
Fig. 12.22. Cell for partial H_3O^+ discharge current (i_H) measurements. A, Anodic compartment; B, cathodic compartment; C, capillary (diameter 0.3 mm); W, water jacket; Fe, iron test electrodes; E, reference electrode; X, T, Z, Teflon taps; G, taps for connection with Nichrome furnaces. (Reprinted from J. O'M. Bockris, A. Despic, and D. M. Drazic, Electrochim Acta **32**: p. 329, copyright 1961, Fig. 1 with permission from Elsevier Science).
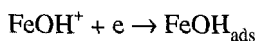
where $K = K_1/K_W$.

With $\beta = 1/2$,

$$\frac{\partial \ln i}{\partial V} = \frac{\partial \ln i}{\partial \eta} = \frac{3F}{2RT}$$

The Tafel slope for the anodic dissolution reaction, b_{anodic} is $\partial \eta / \partial \ln i = 2RT/3F$.

It is easy to show by a similar argument that the rate-determining step for the cathodic direction:

$$FeOH^+ + e \rightarrow FeOH_{ads}$$

yields

$$\frac{\partial \ln i}{\partial \eta} = \frac{F}{2RT} \; ; b_{\text{cathodic}} = \frac{2RT}{F}$$

Equation (12.42) yields

$$\left(\frac{\partial \ln i}{\partial \ln \alpha_{OH^-}} \right)_N = 1$$
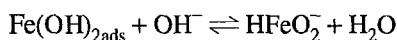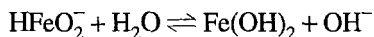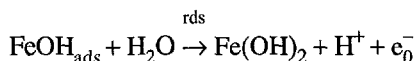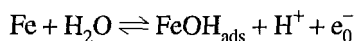
The cathodic kinetics show reaction orders

$$\left(\frac{\partial \ln i_{\text{cathodic}}}{\partial \ln \alpha_{OH^-}} \right)_{N_1 \alpha_{Fe^{2+}}} = 1$$

$$\left(\frac{\partial \ln i_{\text{cathodic}}}{\partial \ln \alpha_{Fe^{2+}}} \right)_{N_1 \alpha_{OH^-}}$$

These kinetic results agree with the facts for $0 < pH < 4$ which, together with the mechanism, were first published by Bockris, Drazic, and Despic (1961).

The mechanism given for dissolution in acid solution applies also in neutral solutions, but with the modification that θ_{FeOH} (assumed to correspond to the condition $\theta < 1$ above) changes to $\theta \simeq 1$.

In alkaline solutions, the mechanism changes. Modern views on it have developed from those of Hackerman (1962) to the present position (Drazic and Hao, 1992):

$$Fe + H_2O \rightleftharpoons FeOH_{ads} + H^+ + e_0^-$$

$$\overset{\text{rds}}{FeOH_{ads} + H_2O \rightarrow Fe(OH)_2 + H^+ + e_0^-}$$

$$HFeO_2^- + H_2O \rightleftharpoons Fe(OH)_2 + OH^-$$

$$Fe(OH)_{2ads} + OH^- \rightleftharpoons HFeO_2^- + H_2O$$

Thus, the mechanism bears a passing resemblance to that at lower pH's but is more influenced by oxides and hydroxides on the electrode surface.

There is no doubt that the all-important reaction of Fe dissolution is unduly sensitive to the precise conditions of the anodic dissolution. The effect of stress has been cited above. But time of measurement and anions present (particularly Cl^-) have a marked effect on the degree of coverage with FeOH and related hydroxides and the

resulting rate-determining step. In future work, spectroscopic and STM examination of the electrode surface while the reactions occur under various conditions will sharpen knowledge in this area. The results available in the 1990s can be put together (Fig. 12.23) in a comprehensive scheme (Drazic, 1989).

12.1.15. The Mechanism of Hydrogen Evolution (HER) on Iron (A Cathodic Partner Reaction in Corrosion often Met in Acid Solution)

Much earlier in this book (Section 7.10), the point was made that the path and rds for hydrogen evolution divide themselves into those (e.g., $H_3O^+ \overset{rds}{\rightarrow} H_{ads} + H_2O$; $2H_{ads} - H_2$) in which θ_H is << 1; and those (e.g., $H_3O^+ + e \rightleftharpoons H_2O + H_{ads}$; $H_3O^+ + H_{ads} + e \rightarrow H_2O + H_2$) in which $\theta \rightarrow 1$. The distinction between these mechanisms is important for Fe because of the effect of the value of the coverage with θ on embrittlement and the related stress corrosion cracking (Section 12.6.5). This is more likely if θ is larger than when it is small because a large θ increases the driving force for the permeation of H into the metal.

For most transition metals, the mechanism of the h.e.r. in acid solution turns out to be the electrochemical desorption step $H_3O^+ + H_{ads} + e \overset{rds}{\rightarrow} H_2O + H_2$, which corresponds to a θ_H close to 1 (embrittling). There are several methods of examining the h.e.r., e.g., the FTIR determination of the H coverage of a surface while evolution occurs (Fig. 12.24), which suggests that if the overpotential is less negative than 0.3 to 0.4 V, the rds is proton discharge, followed by chemical recombination of adsorbed H. At more negative overpotentials, studies of overpotential decay show that the rds remains proton discharge, but that the desorption step changes (which one might expect as θ_H increases) and turns into electrochemical desorption, $H_3O^+ + MH \rightarrow H_2 + H_2O$ (whereupon, because $\theta_H \rightarrow 1$ in this rds, the tendency toward H embrittlement increases).

This trend for a rate-determining proton discharge followed by a chemical desorption step at low η's and an electrochemical one at higher η's seems to survive the change to alkaline solution shown here; of course, the proton discharge occurs from water:

$$H_2O + e \rightarrow H_{ads} + OH^-$$
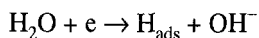
There is a human tendency among the research workers in this field to restrict their work to pure Fe for these studies because they fear complexities caused by alloy inclusions in steel. However, it is steel that is used in practice. Studies to date on steel (Frankenthal, 1986) confirm again a rate-determining proton discharge at lower current densities. However, steel surfaces are heterogeneous and it is likely that some parts of the surface (some crystal planes) are much faster than others.
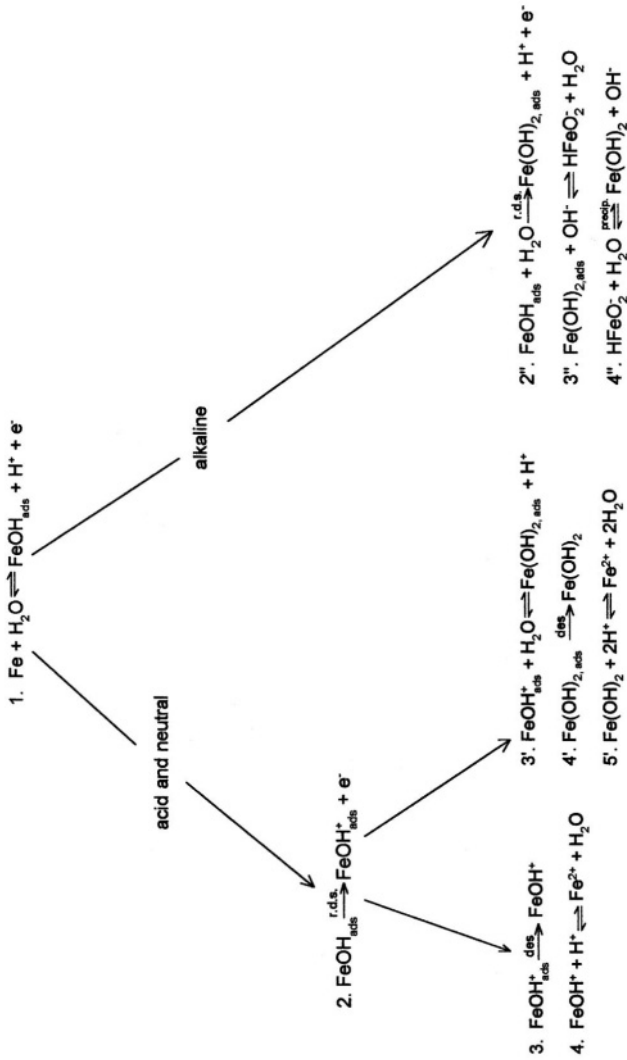
Fig. 12.23. General reaction scheme for iron dissolution (Redrawn from D. M. Drazic, in *Modern Aspects of Electrochemistry*, No. 19, Plenum, 1989).
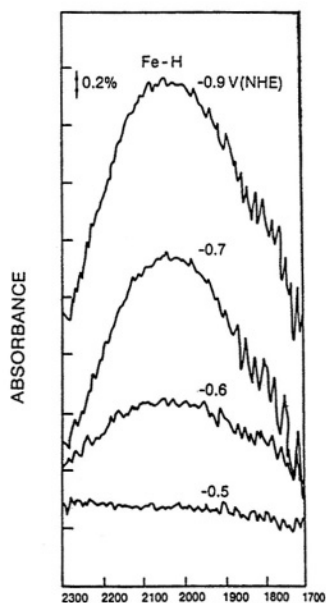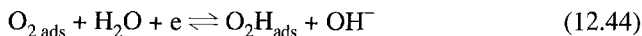
Fig. 12.24. Differential IR spectra in absorbance of Fe-H vibration at different potentials in H_2O borate buffer solution. (Reprinted from J. O'M. Bockris and J. Corbajal, *J. Electrochem*, **134**: 1958, 1987, Fig. 2. Reproduced by permission of The Electrochemical Society, Inc.).

12.1.16. The Mechanism of Oxygen Reduction on Iron

Knowledge here is largely restricted to neutral solutions. It turns out (Jovancicevic, 1986) that the mechanism that best fits the facts for passive iron (see Table 7.15) is

$$O_2 \rightarrow O_{2\,ads} \qquad (12.43)$$

$$O_{2\,ads} + H_2O + e \rightleftharpoons O_2H_{ads} + OH^- \qquad (12.44)$$

$$O_2H_{ads} + e \rightleftharpoons O_2H^- \qquad (12.45)$$

$$O_2H^- + H_2O \rightleftharpoons H_2O_2 + OH^- \qquad (12.46)$$

where O_2H_{ads} follows a Temkin isotherm. The adsorption of the intermediate in the step after the rds follows a Temkin isotherm, so that one writes

$$\gamma\theta/RT = \log[OH^-] + FV/RT + \text{const} \qquad (12.47)$$

$$i = kp_{O_2} \exp(-\alpha FV/RT) \exp(-\beta\Delta G_{ads}/RT) \qquad (12.48)$$

where k is a rate constant, and ΔG_{ads} is the standard free energy of adsorption of the intermediate following the rds, whose free energy of adsorption varies as a function of coverage, θ. Then, with

$$\Delta G_{ads} = \Delta G_{ads}^\circ + \gamma\theta \qquad (12.49)$$

where ΔG_{ads}° is the zero-coverage value, one obtains

$$i = kp_{O_2} \exp(-\alpha FV/RT) \exp(-\beta\Delta G_{ads}^\circ/RT) \exp(\log[OH^-]^{-\beta} - \beta FV/RT) \quad (12.50)$$

Equation (12.50) gives

$$i = k'p_{O_2}[OH^-]^{-1/2} \exp[-(\alpha + 1/2)VF/RT] \qquad (12.51)$$

where k' is another constant and $\beta = 1/2$. This gives an order of reaction with respect to OH^- of -0.5 and the correct O_2 reaction order. The Tafel slope under Temkin conditions for $\alpha = 0$ (chemical rate-determining step preceding an equilibrium electron transfer) gives the correct value of $2RT/F$.

12.1.17. Where We Are Now: Looking Back at the Beginning

The electrochemistry of corrosion is a big piece of electrochemistry. It permeates most of the surface aspects of materials science, at least for practical metal systems in contact with moist air. It influences not only the surface but often the bulk owing to its influence an embrittlement and stress corrosion cracking. So, at the beginning, we argued that a corroding metal is rather like a local fuel cell in which the corroding metal has a very large number of pairs of microsized electrodes on its surface, an equal number of them anodic and cathodic, respectively.

What energy drives an overall corrosion reaction, e.g., $Fe + 2H^+ \rightarrow Fe^{2+} + H_2$? It is, of course, the free energy of the overall corrosion reaction. It must be negative in value for the direction indicated, for otherwise thermodynamics forbids it to occur. This negative free energy of the overall reaction can be converted into the corresponding cell potential, $\Delta G = -nFE$ (n is 2 in the above reaction because two electrons are involved in the constituent electrochemical reaction, $Fe \rightarrow Fe^{2+} + 2e$, and $2H^+ + 2e \rightarrow$

H_2, which make it up). This E of the corroding cell is clearly pH dependent (for H^+ takes part in the reaction) and hence there is a pH dependence of the tendency toward corrosion.

By applying to electrochemical reaction the argument that a positive ΔG means that a reaction having this characteristic cannot occur, Pourbaix and his colleagues have produced many studies of the potentials for metal dissolution, hydrogen evolution, and oxygen reduction in solution of various pH's. Diagrams of the thermodynamically reversible values of these reactions as a function of pH are useful; they tell for certain the conditions under which corrosion cannot occur, i.e., the safe regions of potential for a given pH.

The less well known (so-called) Evans diagrams involve kinetics and allow one to make a rough first cut at the order of magnitude of a corrosion rate (for a pure surface without blocking oxide films).

One of the central fundamental topics in the electrochemistry of corrosion is the atomic-scale mechanism of the sequence of steps by which Fe (the most important metal in construction engineering) dissolves anodically. There are several mechanisms, and the one most investigated has a rate-determining step $FeOH \rightarrow FeOH^+ + e_0^-$. The involvement of OH implies that the anodic dissolution rate is pH dependent.

The electrodics of corrosion reactions yield an approximate equation for the corrosion rate:

$$i_{corr} = \left(i_{0,anodic} i_{0,cathodic} \right)^{1/2} e^{F(E_{cathodic} - E_{anodic}) \frac{F}{4RT}} \qquad (12.32)$$

One can see from this how both kinetics factors (the i_0's) and thermodynamic ones, the E's, affect the corrosion rate. The $E_{cathodic} - E_{anodic}$ tends to be a positive number so that the thermodynamic effects are very large. Finally, a number of practical examples of corrosion are given which show the two reactions that make up a corrosion saturation and just where they occur.

12.1.18. Some Common Examples of Corrosion

This section presents the electrodic principles underlying some familiar instances of corrosion. Automobiles are painted to protect them from corrosion, but often there may be small regions where the steel is exposed to the atmosphere because the paint has been chipped or scratched off. One might at first expect it to be the unprotected metal beneath the broken-paint spot that corrodes. In fact, it turns out that the exposed metal is not the electron-sink area where metal dissolution, or corrosion, occurs. The exposed metal has better access to oxygen than the metal still covered with paint and is therefore the electron-source area; it is the adjoining metal underneath the paint coating that corrodes. *The situation is much worse than it may appear* (Fig. 12.25). A break in paint coatings leads therefore to a spreading of the corroded area rather than to a restriction of corrosion to the exposed spot.
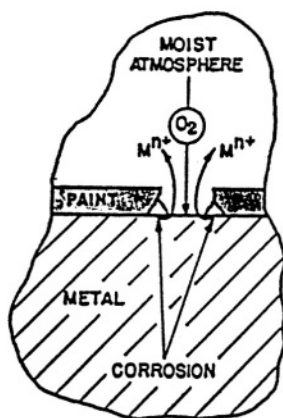
Fig. 12.25. As the paint coating is damaged, the corrosion couple is established, with the metal dissolving at the edges *underneath* the coating, while the exposed part is the place for the electronation of oxygen (and thus not that which dissolves).

This example of corrosion provides a general reason why paint and coatings of various kinds are only a partial answer to the corrosion-prevention problem. They often tend to develop cracks or pinholes, and these exposed spots provide access to oxygen for the electronation reaction, which results in unseen corrosion of the surrounding areas.

The above example of corrosion brings out an important consequence of oxygen reduction as the electronation reaction. This consequence is known as the *principle of differential aeration.* The principle can be stated as follows: If a part of a metal surface has greater access to oxygen, i.e., is in contact with a higher oxygen concentration, than an oxygen-starved area, then oxygen electronation tends to occur at the oxygen-rich area and metal dissolution tends to occur at the oxygen-poor area (Fig. 12.26). In other words, oxygen-rich areas act as electron sources (cathodes); and oxygen-starved regions, as electron sinks (anodes). Thus, owing to the differential accessibility of various parts of a metal surface to the diffusion of oxygen, a corrosion cell is produced with spatially separated electron-source and -sink areas. The exclusion of air (oxygen) from any particular part of a metal system leads to a localized attack of the metal precisely in the oxygen-starved regions.
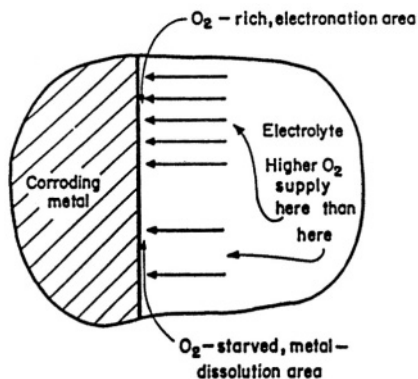
Fig. 12.26. Since the availability of oxygen varies on a metal, metal dissolution takes place at oxygen-starved areas.

One can provide several practical examples of localized corrosion occurring by differential aeration. Crevice attack is a common phenomenon (Fig. 12.27), or, one may mention the corrosion of partially immersed metals in sea water (Fig. 12.28). The region near the waterline provides easy access to oxygen and thus becomes an electron-source area for the lower part of the metal, which becomes an electron sink because of its relative oxygen starvation.

A similar situation prevails when a strip of iron is partly embedded in moist sand underneath water (Fig. 12.29). Contrary to more naive expectations that the sand-covered metal would be protected, it is just this part of the metal strip that dissolves because of its relative oxygen starvation.
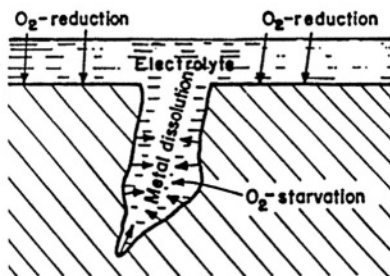
Fig. 12.27. A crevice in the metal surface is a region of limited aeration. Hence, it has a tendency to corrode further and expand.
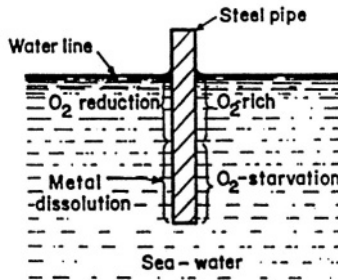
Fig. 12.28. The concentration of oxygen decreases with depth, and this favors metal dissolution in regions farther away from the waterline.

The differential-aeration principle can also be exemplified by the underground corrosion of an iron pipe that runs partly through sand with high oxygen permeability and partly through clay soil with low oxygen permeability (Fig. 12.30); the portion of the pipe in the clay corrodes more heavily than the portion in the sand.

Another example of a differential-aeration corrosion cell is an iron sheet with a drop of moisture on it (Fig. 12.31). The central region of the drop is oxygen starved compared with the peripheral regions, which therefore become electron-source areas, and corrosion is observed at the central electron-sink section.
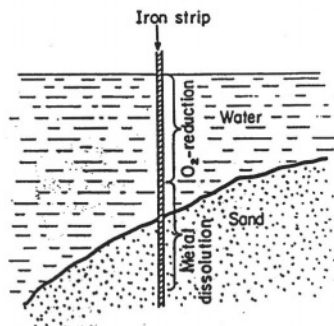

Fig. 12.29. When iron strips are imbedded in sand, it is the imbedded part that undergoes intensive dissolution.
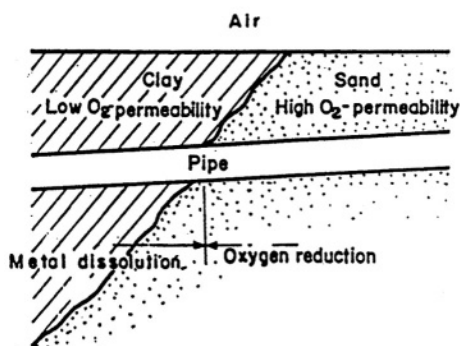
Fig. 12.30. A pipe passing through sand and clay will undergo more intensive corrosion in clay since this has a lower permeability to oxygen than sand.

Apart from corrosion due to differential aeration, corrosion of underground metal structures and pipelines may also arise from stray currents. How this comes about can be seen in the accompanying diagram (Fig. 12.32). The presence of a current-carrying cable in conducting soil results in stray currents passing through the soil. These stray currents may set up a potential difference between two portions of a pipeline, which then develops electron-source (cathodic) and -sink (anodic) areas. Thus, pipelines tend to corrode when they pass near electric lines.
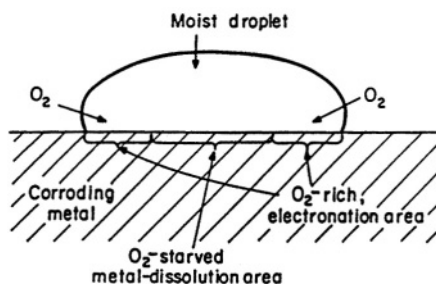

Fig. 12.31. When a droplet of moisture condenses at a metal surface, its center tends to become the point of intensive dissolution activity. The potential distribution associated with the drop was directly measured for the first time by Chen and Mansfeld (1997).
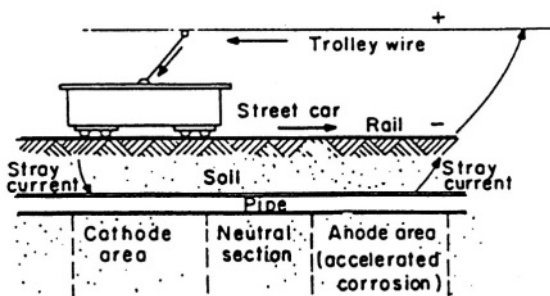
Fig. 12.32. If an underground metal construction comes into a field of stray currents from bare electric lines (e.g., rails), the potential difference established along the construction may set up a corrosion process.

Further Reading

Seminal

1. A. De La Rive, *Ann. Chem. Phys.* **43**: 423 (1830). First suggestion that corrosion had an electrochemical mechanism.
2. L. Cailletet, *Compt. Rend.* **58**: 327 (1864). First report of H embrittlement of metals.
3. A. Finkelstein, *Z. Physikal. Chem.* **39**: 91 (1902). Impedance of passive iron.
4. U. R. Evans, *J. Inst. Metals* **30**: 263 (1923). Evidence in favor of an electrochemical mechanism of corrosion.
5. G. Freenkel and H. Heinz, *Z. Anorg. Chem.* **133**: 167 (1924). The introduction of the term "mixed potential" (a potential determined by two or more individual reactions).
6. C. Wagner and W. Traud, *Z. Elektrochem.* **44**: 391 (1938). The original formulation of the mixed potential concept and the basic theory of corrosion of a pure metal.
7. H. Tomachov, *Dokl. Akad. Nauk. URSS* **30**: 621 (1941). Graphical methods for mixed potentials with many components.
8. M. Pourbaix, "Graphic representation of the relation of pH and Potential in Corrosion" thesis, Delft, University of Technology, the Netherlands, 1945. The original publication of Pourbaix diagrams.
9. B. Kabanov, R. Burshstein, and A. Frumkin, *Disc. Faraday Soc.* **1** :259 (1947). First suggestion of a mechanism of Fe in alkaline solution that was compatible with modern ideas.
10. M. Pourbaix, in *Proc. of the 2nd Meeting of C.I.T.C.E.* (forerunner of *Acta Electrochim.*), Tambarini, Milan (1950). Two fundamental measurements on corrosion.
11. J. O'M. Bockris, in *Modern Aspects of Electrochemistry,* J. O'M. Bockris, ed., Vol. 1, Ch. 4, Butterworths, London (1954). First formulation of equations for the corrosion potential and rate of corrosion in terms of exchange-current densities of the constituent reactions.

12. J. M. Kolatyrkin, *Z. Elektrochem.* **62**: 664 (1958). Alloy corrosion and its mechanism.

13. H. F. Finley and N. Hackerman, *J. Electrochem. Soc.* **107**: 259 (1960). Inhibitors have specific *chemical* effects.

14. J. O'M. Bockris, D. Drazic, and A. Despic, *Electrochim Acta* **4**: 325 (1961). First determination of the cathodic Fe^{2+} deposition current by significantly accurate measurements of the co-evolved H_2; mechanism of the corrosion of iron in acid solution.

15. T. P. Hoar, in *Modern Aspects of Electrochemistry,* J. O'M. Bockris and B. G. Conway, eds., Vol. 3, p. 1, Plenum, New York (1963). On the anodic reactions of metals.

16. E. McCafferty and N. Hackerman, *J. Electrochem. Soc.* **119**: 999 (1972). Mechanism of iron dissolution in the presence of Cl^-.

Modern

1. J. Van Muylder, "Thermodynamics of Corrosion," in *Comprehensive Treatise of Electrochemistry,* J. O'M. Bockris, B. E. Conway, E. Yeager, and R. E. White, eds., Vol. 4, Ch. 1, Plenum, New York (1984).

2. W. H. Smyrl, "Electrochemistry of Corrosion," in *Comprehensive Treatise of Electrochemistry,* J. O'M. Bockris, B. E. Conway, E. Yeager, and R. E. White, eds., Vol. 4, Ch. 2, Plenum, New York (1984).

3. D. Drazic and V. Vassic, *J. Electroanal. Chem.* **155**: 229 (1985). Theoretical analysis of the electrochemical corrosion rate measurement.

4. D. Drazic, in *Modern Aspects of Electrochemistry,* R. E. White, B. E. Conway, and J. O'M. Bockris, eds., Vol. 19, Ch. 4, Plenum, New York (1990). The mechanism of dissolution of iron.

5. A. R. Despic, D. M. Drazic, J. Balaksina, and L. Gejic, *Elektrochim Acta* **35**: 1947 (1990). Mechanism of the dissolution of aluminum.

6. Z. Nagy and R. F. Hawkins, *J. Electrochem. Soc.* **138**: 1047 (1991). Analysis of the correction of the corrosion measurement kinetics for double-layer effects.

7. H. W. Pickering, *Mat. Sci. Eng.* **A198**: 213 (1995). The effect of ohmic drop on corrosion measurements.

8. I. H. Plonski, in *Modern Aspects of Electrochemistry*, J. O'M. Bockris, B. E. Conway, and R. E. White, eds., Vol. 29, Ch. 3, Plenum, New York (1996). Effect of adsorbed H on Fe dissolution rate.

9. J. O. Park, C. H. Park, and R. C. Alkire, *J. Electrochem. Soc.* **145**: L174 (1996). Measurements of corrosion in small species.

10. C. Wang, S. Chen, and X. Yu, *J. Electrochem. Soc.* **143**: L283 (1996). Anodic dissolution of iron on a magnetic field with holographic microphotography.

11. R. Wagner, *J. Electrochem. Soc.* **193**: L139 (1996). Copper corrosion in thin films of acid.

12. C. C. Chen and F. Mansfeld, *Corros. Sci.* **39**: 409 (1997). Potential profile under drop of solution on steel.

13. J. O'M. Bockris and Y. Kang, "The Mechanism of the Corrosion of Al Alloys," *J. Solid State Electrochem.* **1**: 17 (1997).

12.2. INHIBITING CORROSION

12.2.1. Introduction

Looked at from afar, one can see that the corrosion of metals bears some similarity to the aging of biological systems. Now, just as, in recent years, it is being claimed that biological aging can be retarded by consuming dietary supplements such as vitamin E and dehydroepiandrostenone (DHEA), so there are several ways including the addition of organic substances to a solution in contact with the metal, to retard (in some cases, to a high degree) the spontaneous electrochemical dissolution known as corrosion.

Two main approaches to this important subject exist. If the object (as with a ship's hull) is in contact with an unlimited amount of aqueous solution, addition of a chemical to the "solution" is not feasible. For this kind of situation there are two electrochemical approaches: cathodic and anodic protection. However, often (as with oil pipelines) if the corroding liquid (e.g., sea water) is at least partially confined,[4] then there is great value in developing organic molecules that adsorb on the metal and reduce the velocity of anodic dissolution.

Before launching into a description of these methods, it is salutary to make a statement about the cost of the corrosion taking place around us and open for electrochemists' initiative. Although such estimates change with time and place, many informed calculations for the United States and United Kingdom have hovered between 1 and 2% of the gross national product. It would be difficult to find a topic in technology in which the wide application of (even) present knowledge, would yield such a great financial gain.

12.2.2. Cathodic and Anodic Protection

Let it be reiterated that the control and prevention of corrosion is a subject with tremendous technological and economic significance. It is not a surprise, therefore, that the literature on this subject contains vast amounts of empirical information. No attempt is made here to survey these details. Rather, it is intended to present in a simplified way the essence of the electrochemical approaches to corrosion prevention.

A good starting point is the basic picture of corrosion by local-cell action, according to which a corroding metal consists of electron-sink areas at which metal dissolution takes place and electron-source areas at which an electronation reaction occurs. Under conditions where there is an exponential current-potential relationship

---

[4]When oil from undersea sources flows through a pipeline, it is sometimes accompanied by a saline water layer that sequesters itself between the flowing oil and the metal or metal oxide surface of the pipe. The sea water does not remain stationary and much of it eventually exits with the oil (after which it gets separated and ejected to the sea). Nevertheless, it remains in contact with pipe material for the time of its journey from the beginning of the pipe below the seabed until its exit on the platform.

for both the metal-dissolution and electronation reactions, and the transfer coefficients are equal to $\frac{1}{2}$, the corrosion current has been shown to be given by

$$I_{\text{corr}} = (I_{0,\text{M}}I_{0,\text{SO}})^{\frac{1}{2}} \exp \frac{F(E_{0,\text{SO}} - E_{e,\text{M}})}{4RT} \tag{12.32}$$

Two fundamental ways in which the magnitude of I_{corr} can be reduced may be seen from (12.32). The first method is based on diminishing the product $I_{0,\text{M}}I_{0_{\text{SO}}}$; this is the method called *corrosion inhibition*. The second method is based on making the relative potential of the corroding metal, E, equal to or less than the equilibrium potential $E_{e,\text{M}}$ for the metal-dissolution reaction; this is the method called *cathodic protection*. The basic principles of these two methods of corrosion control and prevention will now be presented.

12.2.2.1. Corrosion Inhibition by the Addition of Substances to the Electrolytic Environment of a Corroding Metal. Consider the ways in which the term $I_{0,\text{M}}I_{0,\text{SO}} = (A_{\text{M}}i_{0,\text{M}})(A_{\text{SO}}i_{0,\text{SO}})$ can be reduced. First, one can try to reduce the exchange current densities of the metal-dissolution and electronation reactions. For instance, if hydrogen evolution is the electronation reaction, the addition of phosphorus, arsenic, or antimony compounds (e.g., As_2O_3) reduces the exchange current density for hydrogen evolution (Fig. 12.33). Or, if oxygen electronation is the reaction at the electron-source areas, then, by adding substances that react with dissolved oxygen, the O_2 concentration is reduced and therefore, also, the exchange current density for oxygen reduction. Two substances that act in this way are hydrazine, N_2H_4, and sulfite ions, SO_3^{2-}

$$2N_2H_4 + 5O_2 \rightarrow 4NO_2^- + 4H^+ + 2H_2O \tag{12.52}$$

and

$$2SO_3^{2-} + O_2 \rightarrow 2SO_4^{2-} \tag{12.53}$$
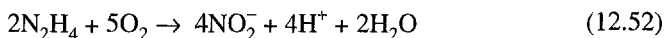
Alternatively, or in addition, the exchange current density for the metal-dissolution reaction can be reduced by adding compounds that adsorb on the electron-sink areas of the corroding metal and slow down the metal-dissolution reaction (Fig. 12.34). The compounds most often used for this purpose are nitrogen-containing organic compounds (aliphatic and aromatic amines), sulfur-containing compounds (thiourea and its derivatives), and various oxygen-containing compounds (aldehydes).

Now, it will be recalled that the adsorption of a particular constituent of the electrolyte depends not only on its chemical nature (i.e., on the chemical part, ΔG_c°, of its free energy of adsorption) but also upon the electrode charge (remember the parabolic θ_{org} versus q_{M} curves in Chapter 7). So the corrosion inhibitor must not only be highly adsorbable in a chemical sense, it must also adsorb in the range of potentials which includes the potential at which the corrosion reactions occur. Correspondingly,
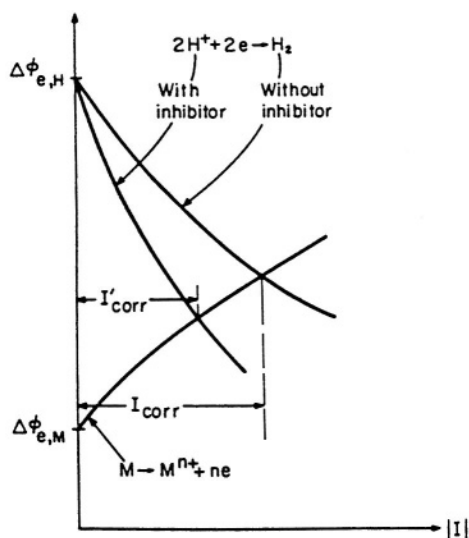
Fig. 12.33. If a hydrogen-evolution reaction is inhibited, the potential–current relation changes in such a way that the corrosion current is reduced (i.e., I_{corr} is reduced to I'_{corr}).

differences in the degree of coverage arise when a metal is polarized cathodically or anodically in respect to the corrosion potential.

A second way of reducing the product $(A_M i_{0,M})(A_{SO} i_{0,SO})$ is to reduce the areas of the corroding metal that function as electron sinks or sources. The reduction of the electron-sink area is usually achieved by means of the solid products of metal dissolution. This, however, is a process that will be discussed later on (see Section 12.5). Here, reference will be made to methods of producing solid films that reduce the electron-source areas. What is done is to add a film-forming inhibitor that causes the precipitation of a solid film over the electron-source areas. An example of a film-forming inhibitor is the HCO_3^- ion, which interacts with the OH^- ions proeduced by oxygen reduction to precipitate a carbonate film over the electron-source areas. Another example is the PO_4^{3-} ion, which causes the precipitation of a mixture of ferrous and ferric phosphates over steel.

Film-forming inhibitors that affect the electron-source areas may be distinguished from those that affect the electron-sink areas by the fact that the former alter the corrosion potential in the negative direction (Fig. 12.35), and the latter, in the positive direction (Fig. 12.36). However, although an inhibitor may start its action on the electron-source areas, it may continue causing the precipitation over the whole surface of the corroding metal and produce a general blockage of the metal surface. With
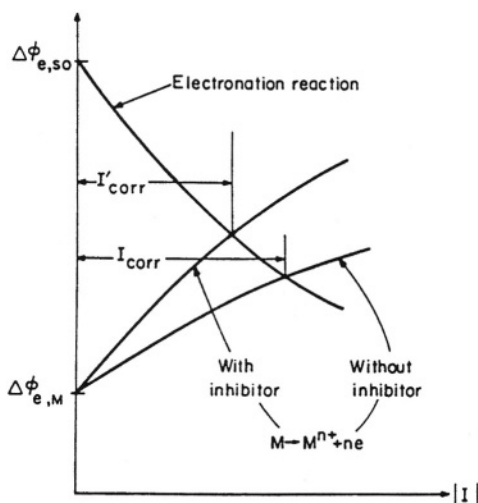
Fig. 12.34. A metal dissolution reaction can be inhibited by adsorbing organics at electron sink areas, in which case the corrosion current is reduced.

inhibitors that produce a general coverage of the surface, the corrosion potential may move either way, depending on which reaction is affected more.

12.2.2.2. Corrosion Prevention by Charging the Corroding Metal with Electrons from an External Source. A metal corrodes because the potential difference across the interface at the electron-sink areas is positive with respect to the equilibrium potential for the metal-dissolution reaction. If by some means the potential difference could be made negative with respect to the equilibrium potential, metal dissolution would not occur. This depression of the potential difference between a metal and its corrosive environment can be achieved by arranging for electrons to be pumped into the corroding metal. These electrons will make the metal more negatively charged and thus lower the potential difference in the negative direction. To prevent dissolution of the metal, it is necessary to pump in an adequate number of electrons.

One method of pumping electrons into the corrodible metal is based on a well-known electrodic fact. When the ions inside a suitably selected metal pass into solution, they leave behind excess electrons, which, if provided with an electronically conducting path, can be made to flow into the corrodible metal. Suppose that an auxiliary metal having an equilibrium potential negative to that of the corrodible metal is immersed in the corrosive environment and connected by a short-circuiting wire to the metal to be protected (Fig. 12.37). Then the auxiliary metal will function as an electron sink (anode) and sacrificially dissolve (hence the term *sacrificial anode*).
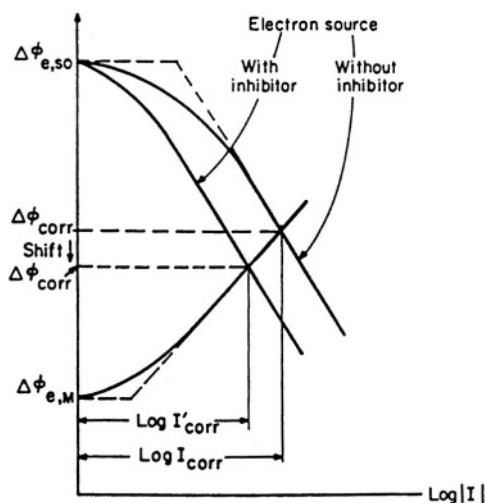
Fig. 12.35. If an inhibitor adsorbs on electron source areas, the corrosion potential shifts in the negative direction.

Further, the corrodible metal will act as the electron-source electrode for the electronation reaction, which would otherwise have produced its corrosion. Hence, what is done is to set up a new corrosion cell in which an auxiliary metal is made to corrode in place of the metal to be protected and in which the entire surface of the latter metal is converted into an electron-source area. For example, if a steel structure has to be protected, one can use zinc or magnesium as a sacrificial electron sink and save the structure from corrosion.

The electrons pumped into the corrodible metal have come, in the above method, from the dissolution of a scarificial auxiliary metal. Instead, they can come from an external current source (i.e., an electrical power supply). The electrical circuit, however, has to be completed, and toward this end, an auxiliary *inert* electrode can be immersed in the corrosive electrolyte to provide a return path for the electron current (Fig. 12.38). The external source can then be adjusted so that the potential difference between the corrodible metal and its environment becomes negative with respect to its equilibrium potential. Under these circumstances, the whole of the metal to be protected against corrosion will function as an electron source for the electronation reaction, and the second electrode will serve as an electron sink for some deelectronation reaction (Hoar).

The above two methods of preventing corrosion can be understood easily with an Evans diagram (Fig. 12.39; see Section 12.19). (These diagrams it will be recalled, result from the superposition of the potential–current curves of the electronation and
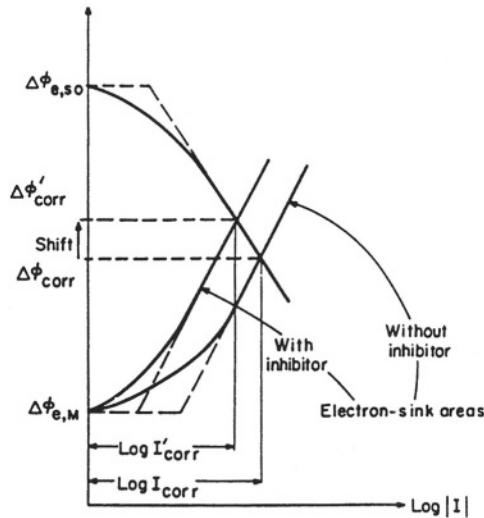
Fig. 12.36. If an inhibitor adsorbs on electron sink areas, the corrosion potential shifts in the positive direction.

deelectronation reactions that occur during corrosion.) In spontaneous corrosion reactions, the electronation current is equal to the metal-dissolution current at the corrosion potential. With an external current source for adjusting the potential difference between the corroding metal and its electrolytic environment, it follows that as the metal-solution potential difference becomes more negative, the metal-dissolution current decreases, whereas the electronation current increases, so that
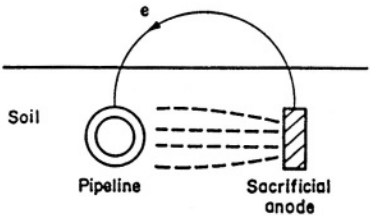
Fig. 12.37. Cathodic protection can be accomplished if an auxiliary metal of a more negative potential is short-circuited with the metal construction to be protected.
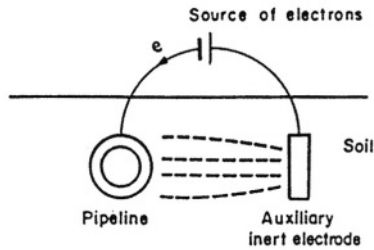
Fig. 12.38. Cathodic protection can also be accomplished by using an inert auxiliary electrode and *sending current into the circuit* from an external current source.
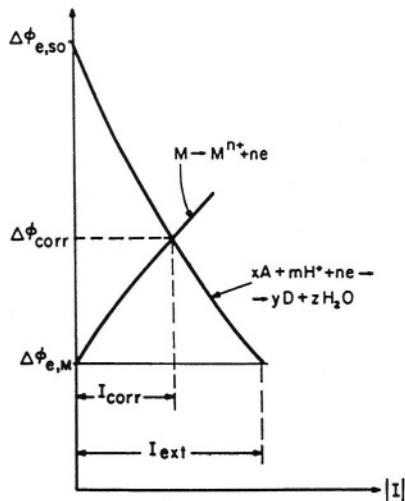
Fig. 12.39. The change in potential from $\Delta\phi_{corr}$ to $\Delta\phi_{e,M}$ *results in a decrease of the metal dissolution current to zero*, but also in an increase in the electronation current, and this increase in current has to be supplied from an external source.

$$I_{SO} > I_M \tag{12.54}$$

When the potential of the previously corroding metal becomes equal to the equilibrium potential for the dissolution reaction,

$$I_M = 0 \quad \text{and} \quad I_{ext} = (I_{SO})_{\Delta\phi_{e,M}} \tag{12.55}$$

where I_{ext} is the current that has to be supplied by the external source to protect the corroding metal from corrosion.

The stabilization of metal surfaces by the superimposition of an adequately negative potential difference across the interface between the metal and its environment appears to be an ideal method of corrosion prevention. There are, however, some less favorable aspects of the method.

First, when external current sources are used, the power consumption may be impractically large. It all depends on the electrodic parameters of the electronation reaction—the larger its exchange-current density and the lower its Tafel slope, the larger will be the external protection current that must be used to achieve protection.

Second, it is important that the potential difference across the *entire* interface between the metal to be protected and its environment be shifted below the equilibrium potential. Suppose that the current through the electrolyte is not distributed uniformly over the corroding metal (e.g., the current lines may pass through greater distances to reach some parts of the metal and hence introduce an IR drop near those parts); there may then be localized areas at which the potential difference remains insufficiently cathodic and metal dissolution occurs. Under these circumstances, one may be deceived into thinking that an adequate protection against corrosion has been set up, whereas in fact localized corrosion is occurring. Also, it is often better (e.g., in a pipeline) to suffer a very slow, uniform dissolution than localized attack and puncturing.

Conversely, under conditions where the surface of the metal is made *excessively* negative with respect to the hydrogen equilibrium potential, excess hydrogen evolution may occur and one has to reckon with a hazardous consequence of such cathodic protection. While the metal is successfully protected from dissolving, its surface is being covered with adsorbed hydrogen atoms that are intermediates in the hydrogen-evolution reaction. It will be shown later that some of those adsorbed hydrogen atoms may dissolve into the metal and that this hydrogen which is thus pumped into the interior of the metal can *undermine* the internal strength of the metal by what is called *hydrogen embrittlement* (Section 12.6.6). One may end up, therefore, by losing the strength of the body of the metal.

12.2.3. Anodic Protection

The above sections outlined two methods for the inhibition of corrosion. One (cathodic protection) is entirely electrochemical in nature and consists in moving the

potential of the sample in the cathodic direction with respect to the corrosion potential. Visualization of the Butler–Volmer equation in a form applied to corrosion rate [e.g., Eq. (12.34)] shows that a shifting of the potential by about 100 mV in the negative direction cuts the corrosion rate by more than 10 times. Of course, the corresponding increase in the cathodic reaction over that at the corrosion potential could possibly cause trouble, namely hydrogen embrittlement (Section 12.6.6) and a consequent reduction in the ductility and strength of the metal.

However, there is another way of reducing corrosion, although it may apply only to certain metals. Thus, if, instead of moving the potential in the *cathodic* direction (and cutting down the fundamental rate of dissolution), one goes in the other direction and makes the potential of the sample more anodic, then another mode of inhibition may be attainable. At first, this seems anomalous. Would not a more anodic potential *increase* the anodic dissolution current? In some of the simpler cases (e.g., Cu in $CuSO_4$ solution) this is what would indeed happen. However, in other cases (and the all-important iron and aluminum corrosion is an example), going more anodic increases the rate of alternative anodic reactions, the discharge of H_2O or OH^- onto the metal to form an oxide film. These films vary very much in character. Some, however, only 10–20 monolayers thick, may be very resistive to transfer through them of metal ions that are solution bound. These films simply shut down to a great extent the anodic dissolution and build M–O bonds on top of the metal to be protected.

A special kind of anodic film formation is called *passivation* and it is itself the best example of anodic protection. So important has passivation been over the years, and so puzzling its mechanism, that the subject (which implicitly presents anodic inhibition) will be treated in a separate section (12.4).

12.2.4. Organic Inhibition: the Fuller Story

It has long been known that the addition of certain substances (inorganic salts, organic molecules) to solutions in which a metal is corroding may substantially reduce the corrosion rate.[5] (See Figs. 12.33 and 12.36). One can link the power of an organic molecule to inhibit corrosion to several electrochemical properties. Each metal–solution system has a "corrosion potential," and it was shown in Section 11.1.9 how this particular potential depends on a mixture of the exchange current densities (hence the rate constants) of the cathodic and anodic reactions making up the corrosion couple,

---

[5]T. P. Hoar, who was co-discoverer with U. R. Evans (in 1936) of the basic facts about the electrochemical mechanism of corrosion that led to Wagner and Traud's seminal theory, told about an episode from his early days as a corrosion consultant. Approached by an automotive concern for an inhibitor to stop the distressing breakdowns of its 1930s car radiators, he busied himself in his lab over a weekend and created (stumbled upon?) a potent organic inhibitor for the system concerned. The client wanted to pay a handsome fee, but Hoar wisely tempered his enthusiasm and humbly asked for just a few cents for every time the inhibitor was used. His decision, he says, provided him with a significant income for more than a decade.

together with the two thermodynamically reversible electrode potentials of these reactions, respectively.

Now, the adsorption of organic molecules (and they must be adsorbed to become inhibitors!) is understood, also, and the simpler organics maximize their adsorption near the potential of zero charge (Section 6.9.3) of the metal or its oxide. This does not mean that if the corrosion potential of the system differs by, say, 0.1 V from the pzc, the organic molecule will no longer inhibit. Thus, in Figs. 12.40 and 12.41, it is seen that the bell-shaped θ–V curve for the adsorption of simpler organics stretches over at least 0.8 V (Jeng, 1993). There are some aromatic inhibitors (those in which the π bonds of the organic to the metal are the cause of much of the heat of adsorption) in which the range of potentials in which the organic is highly adsorbed on the surface is very large (more than 1 V). This means that the rule; "optimal inhibition by organics is at the pzc and inhibition is best if the corrosion potential coincides with this," is a simplistic generalization, but is helpful in deciding (knowing the θ–V relation of an inhibitor) what inhibitors will work for the corrosion of a given metal.[6] It could be, for example, that the pzc and corrosion potential are more than 0.5 V apart and then the inhibitors most likely to be useful will certainly be those having plenty of aromatic character, for then even if there is the difference quoted, the inhibitor is likely to be still highly adsorbed at the corrosion potential and hence effective in inhibiting corrosion.

The sketch given here of the relation of the potential dependence of organic corrosion inhibitors and the corrosion potential is quite schematic and there is much detail to add. In this section alone, three points can be mentioned.

1. Some inhibitors are remarkable in their effectiveness. An excellent illustration of this lies in oil extraction technology in the more difficult situations that increasingly arise as the more accessible oil pools become exhausted and the remaining oil is found only at great depths (over 8000 ft) and sometimes under a thick shield of rock. To get at an oil supply often worth hundreds of millions of dollars, it may be most economic to eat through the rock by means of concentrated hydrochloric acid at or near the boiling point. How may this solution be kept working, week after week, until the rock is dissolved away and the still precious (though polluting) fuel becomes available? The method used is to manufacture a stainless steel tube (more than 1000 m in length!) from the surface leading to the rock face. Boiling HCl is introduced into this tube. The tube itself is worth nearly $1 million. Boiling concentrated HCl corrodes stainless steel at a high rate. However, the situation has been made economically viable by the discovery that 1-octyne-3-ol ("octynol"), dissolved in the boiling mixture, reduces corrosion of the stainless steel pipe by about 99%.

2. The actual orientation of octynol in the position at lower concentrations corresponds to a prone position (see Fig. 12.42). However, at higher concentrations, when the tendency from solution is for a greater number of molecules to be adsorbed

---

[6]An ideal inhibitor would be one in which the mechanism in the θ–V curve coincides with the corrosion potential of the metal to be protected.
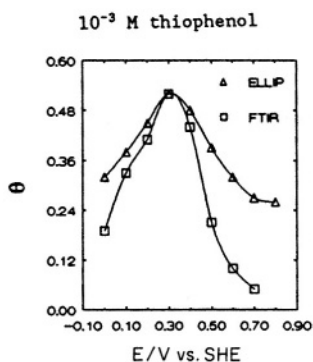
10^{-3} M thiophenol

Fig. 12.40. A relatively small potential range for adsorption will inhibit only if the corrosion potential is $\pm$ 0.2–0.3 from the pzc. (Reprinted from J. O'M. Bockris and K. T. Jeng, *J. Electranal. Chem.* **330**: 541–581, copyright 1992, Fig. 6p, with permission from Elsevier Science.)

per unit area, the molecules are forced to stand up (Fig. 12.43) to make room for the extra ones pressing in from the solution.

    3. It will be seen in Section 12.4 on passivation that protective oxide films break down particularly easily in the presence of Cl^-, and hence Cl^- competition in adsorption with an inhibitor such as octynol takes place. In fact, the presence of Cl^- slows down octynol adsorption.

10^{-6} M 1-naphthoic acid
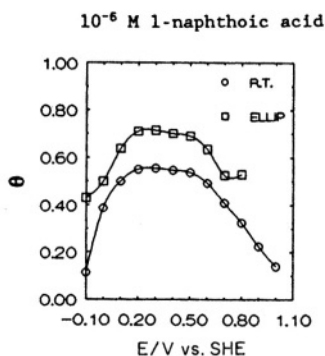
Fig. 12.41. A broader coverage of the potential range is helpful in inhibition because it can match a wide range of corrosion potentials. (Reprinted from J. O'M. Bockris and K. T. Jeng, *J. Electranal. Chem.* **330**: 541–581, copyright 1992, Fig. 6f with permission from Elsevier Science.)
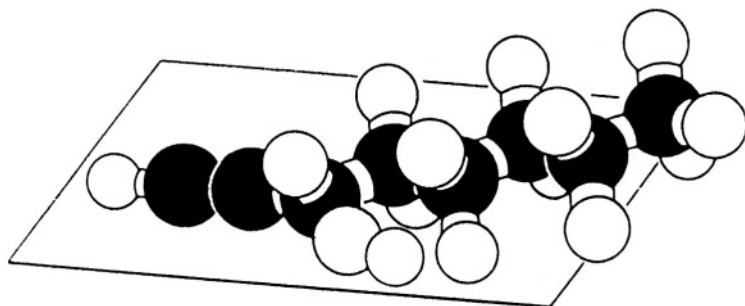
Fig. 12.42. Adsorption of 1-octyne-3-of on iron in a flat orientation. (Reprinted from V. Jovancicevic, B. Yang, and J. O'M. Bockris, "Effects of Chloride on the Adsorption, Orientation, and Polymerization of 1-Octyne-3-of on Iron" in *Electrochimica Acta* **32**:1558, copyright 1987, with permission from Elsevier Science.)
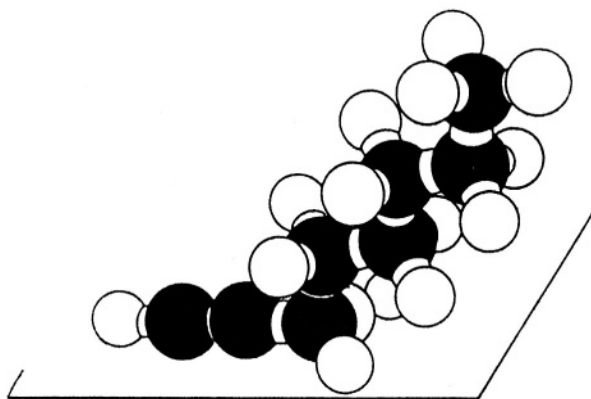
Fig. 12.43. Adsorption of 1-octyne-3-of on iron in a partly upright orientation. (Reprinted from V. Jovancicevic, B. Yang, and J. O'M. Bockris, "Effects of Chloride on the Adsorption, Orientation, and Polymerization of 1-Octyne-3-of on Iron" *Electrachimica Acta* **32**:1558, copyright 1987, with permission from Elsevier Science.)

12.2.5. Relations between the Structure of the Organic Molecule and Its Ability to Inhibit Corrosion

It was indicated in the last section that most organic compounds maximize their adsorption near the potential of zero charge of the metal on which they adsorb. This suggests that organic molecules are more likely to be effective in inhibiting corrosion to the extent that the corrosion potential is near the pzc. However (see Fig. 12.41), the potential range in which the inhibitor adsorbs, around the pzc, can be large, particularly with aromatic compounds.

Another aspect of the adsorption of organics on metals, which was not elaborated on in the theory (6.9.3) that underlies the parabolic dependence of θ_{org} on V, is the strength of the adsorption, its ΔG_{ads}°. This will reflect several energy quantities. Thus, the organic first has to displace the water molecules adsorbed on the metal. The energy for this, per mole of organic, will be proportional to the water–metal bond strength and the number of water molecules displaced when an organic arrives from the solution to adsorb on the metal. Then, clearly, the actual bonding of the organic to the metal will often be dominating and in the 1990s, the possibility of quantum mechanical estimates of such bonding and how it depends on structure was in the foreground (Hackerman, 1995).[7] Finally, the organic has to be taken out of the solution onto the electrode partly desolvated, so that its energy of solvation (and hence the organic's solubility) will be important, as well as any competition it may have with adsorbing anions such as Cl^- (Fig. 12.44). An early experimental relation between adsorbability and solubility and their dependence on structural groups was determined by Blomgren, Jesch, and Bockris (1961). It is shown in Fig. 12.45.

In spite of the advances made in quantum mathematical approaches to the strength of adsorption, the structures that make good inhibitors are still recognized more by empirical rules (which in some sense can be seen from a knowledge of bonding) than by numerical calculation. Thus, the following classes of compounds are those in which corrosion inhibitors are mostly sought.

- Aromatics, phenyl and naphthyl compounds
- Aliphatic unsaturated compounds, e.g., the acetylenic, octynol, as mentioned above
- Alicyclics, such as the pyroles
- Aromatic heterocycles, e.g., quinolines
- Sulfur-containing compounds, e.g., thiourea

All these classes of compounds have provided corrosion inhibitors. Out of this very extensive work (Hackerman, 1989; Singh and Lin, 1997), it is possible to come

---

[7]However, a full quantum mechanical ab initio calculation of bonding of one of the large organics to the metal takes too long for currently available hardware. As shown later, quantum mechanical reasoning can support known qualitative trends (and those discussed in this section) to help design inhibitors (Lin, 1997).
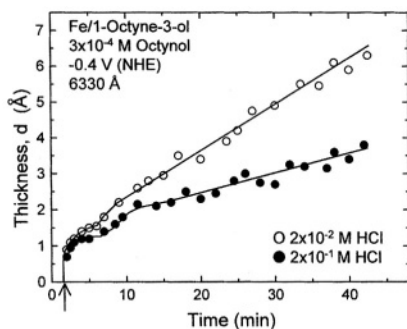
Fig. 12.44. Thickness–time relationship for the adsorption of 1-octyne-3-of on iron for various concentrations of Cl⁻. (Reprinted from V. Jovancicevic, B. Yang, and J. O'M Bockris, "Effects of Chloride on the Adsorption, Orientation, and Polymerization of 1-Octyne-3-of on Iron," *Electrochimica Acta,* **32**: 1558, copyright 1987. Fig. 2, with permission from Elsevier Science.)

to the following generalizations as to trends for elements of structures in an organic that give rise to good inhibition:

- Aromatics and the presence of π bonds
- High electric polarity (dipole moment) of substituent groups in the organics
- Availability of lone pairs in the substituent groups
- Long side chains up to about 11 carbon atoms
- Energy of the highest occupied orbital (HOMO) and its compatibility with the energy of the lowest unoccupied orbital (LUMO) in the metal

A good inhibitor does not have to have all these characteristics, but one at least must be present. Table 12.2 contains examples of excellent inhibitors of the corrosion of iron that were well known in the late 1990s.

The importance of a hydrocarbon chain may be connected with its rotation (Fig. 12.46). The inhibition efficiency depends, of course, on concentration of the inhibitor in solution and hence on its solubility (Fig. 12.47). Another interesting fact is the great variation of the concentration at which saturation of the surface by the inhibitor is reached. Thus Fig. 12.47 shows a variation over three orders of magnitude although the same type of compound (alkyl ammonium salts) comprise the data set. Thus, with
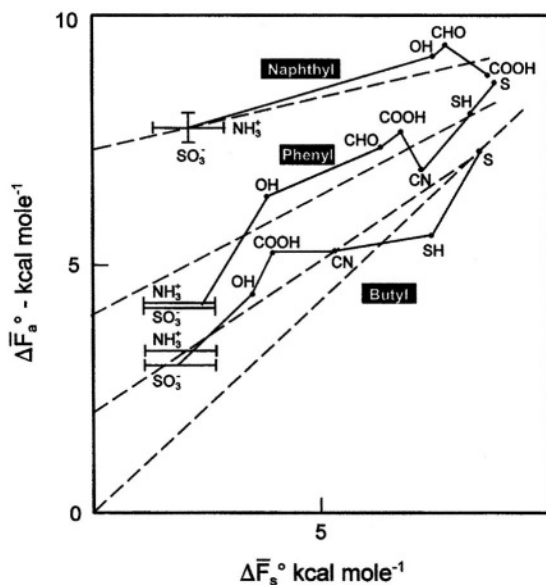
Fig. 12.45. The standard free energy of adsorption of families of organics as a function of the standard free energy of solution. Note the difference in sign. Adsorption occurs more readily as the tendency to dissolve declines. (Reprinted with permission from E. Blomgren, J. O'M Bockris, and K. Jesch, *J. Phys. Chem.* **65**: 2006, copyright 1961 American Chemical Society.)

increasing complexity of the salt, the concentration for saturation can be as low as 10^{-5} M (Fig. 12.48).

12.2.6. Toward a Designer Inhibitor
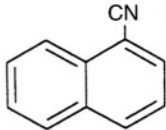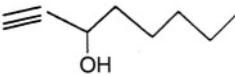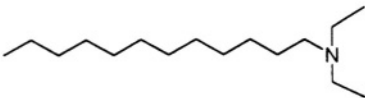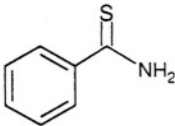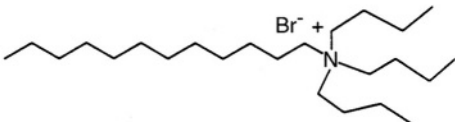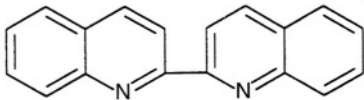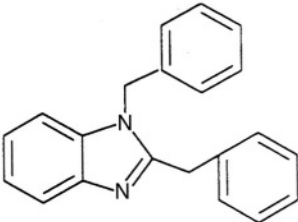
The forward look in the electrochemistry of corrosion inhibitors is the theoretical design of inhibitors. The cutting edge in this field is in such design work (Hackerman, 1995; Singh and Lin, 1997). Consider the theoretical interpretation of the action of an inhibitor, dibenzyl sulfoxide:
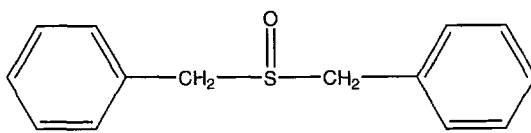
Dibenzyl Sulfoxide

In Chapter 7, the importance of knowing the rate-determining step in an electrode reaction was stressed, particularly when it came to designing electrocatalysts. In

TABLE 12.2
A Few of the Best Inhibitors for Preventing the Corrosion of Iron and Its Alloys

Naphthylcyanide

1 - Octyn - 3 - ol

N, N′ diethyl dodecylamine

Thiobenzamide

N - dodecyl, tributylammonium bromide

2, 2′biquinoline

1,2-dibenzyl benzimidazole

corrosion, two reactions are always involved. Although in the steady state, the rate of the anodic dissolution reaction must be equal in magnitude and opposite in direction to its cathodic partner reaction, one of these two will be the rate-determining one. Cathodic and anodic inhibitors (i.e., those that act in the cathodic reaction, or, alternatively, the anodic, exhibit different characteristics, as shown in Fig. 12.49.

The first thing that has to be done, therefore, is understanding (from the behavior of the results in diagrams such as the above) whether the principal action of the inhibitor should be on the cathodic partner reaction or on the anodic one. For example, if an inhibitor acts cathodically, this implies that before the inhibitor was introduced, the rate-determining step in the corrosion reaction was the cathodic one.
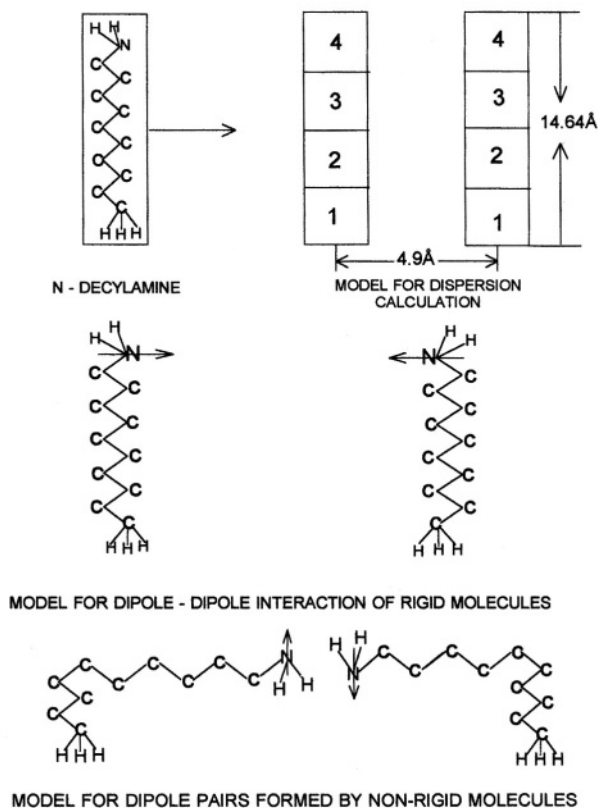
N - DECYLAMINE

MODEL FOR DISPERSION
CALCULATION

MODEL FOR DIPOLE - DIPOLE INTERACTION OF RIGID MOLECULES

MODEL FOR DIPOLE PAIRS FORMED BY NON-RIGID MOLECULES

Fig. 12.46. Models for lateral interaction calculations between *n*-decylamine molecules. (Reprinted from J. O'M Bockris and D. A. J. Saunklec, *J. Electrochem. Soc.* **111**(6): 741, 1964, Fig. 8. Reproduced by permission of the Electrochemical Society.)

For dibenzyl sulfoxide on iron, it turns out that indeed it is the hydrogen evolution reaction, the partner reaction to the anodic dissolution reaction, which controls the corrosion rate. Because the inhibitor acts cathodically, it must interfere with and slow down the rds of this reaction , i.e., make it more difficult for the H to be desorbed.

This is where the techniques of computational chemistry come in. A cluster of 96 iron atoms has been studied (Kutej and Hackerman, 1995) and from this it is possible to derive the charges on the H atom of the surface plane (Fig. 12.50). A corresponding calculation of charges on the surface after adsorption of the organic (assumed to be in reduced form) is shown in Fig. 12.51.
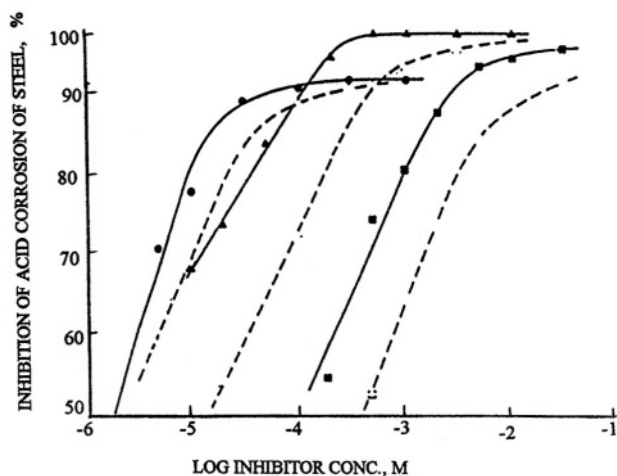
Fig. 12.47. Corrosion inhibition by *n*-alkyl-triethyl- and -trimethylammonium bromides in 1 N-H_2SO_4 at 20 °C. (Reprinted from W. P. Singh, "Relationships between the Structure of organic Compounds and Corrosion Inhibition," in Green Inhibitor Consortium, Texas A&M University, 1997.)

The inhibitor withdraws electrons from the Fe. This has the effect of increasing the bond of the adsorbed H to the metal and reduces the rate of the cathodic evolution of H_2 and therefore that of the partner anodic dissolution, which must function at the same rate as that of hydrogen evolution.

Another computer-oriented approach to inhibitor design (Lin, Singh, and Bockris, 1996) is to use data already available and a software program that examines these data (on corrosion inhibitor efficiency) for correlation with a number of molecular properties belonging to the inhibitor concerned. Some 40 characteristics of each of the molecules concerned are tried out by the computer to see if any one of these properties correlates with the efficiency of corrosion shown by the compounds for which corrosion inhibition data are available. If several molecular properties of the would-be inhibitors are indeed found to correlate with the corrosion inhibition of the "learner" compounds, a "learning graph" can be constructed between the known corrosion inhibitor's efficiency for the learner compounds and an expression that involves the properties that have been found to affect the degree of corrosion inhibition provided by each molecule. With such a graph, it is possible to take any molecule and (knowing the numerical value of its characteristics which correlate with its properties as a corrosion inhibitor) find where it falls on the graph, i.e., determine its efficiency as a corrosion inhibitor.
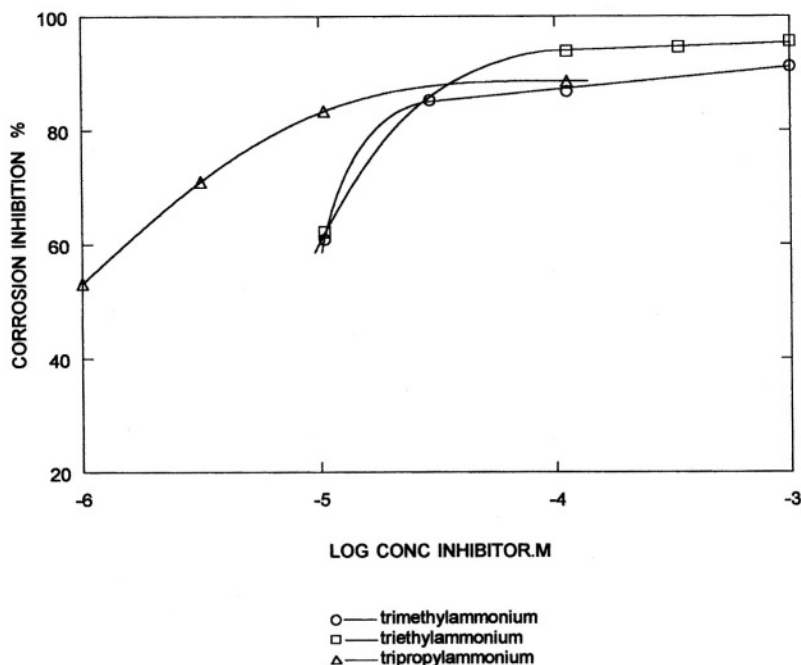
Fig. 12.48. Corrosion inhibition for pure iron in 0.5 M H_2SO_4 containing *n*-hexadecyl quaternary ammonium bromides with different head groups. O-O, trimethylammonium; □-□, triethylammonium; △-△, tripropylammonium.

This method is, of course, a sophisticated empiricism, but it allows the corrosion efficiency of any organic to be rapidly estimated. The reservation is that there must have been some prior data on corrosion inhibition for the system concerned (e.g., Fe in acid solution) to provide the learning graph.

12.2.7. Polymer Films as an Aspect of Corrosion Inhibition

In Section 12.2.4, the remarkable efficacy of octynol in reducing the rate of corrosion of stainless steel by 99%, even in concentrated boiling HCl, was noted. Such a remarkable inhibitory solution has been examined in detail (Bockris and Bo Yang, 1991) using XPS, ellipsometry, and a technique (MacFarlane, 1990) that uses a study of the bombardment of a surface by particles emitted by the new element, Californian, to obtain information on groups adsorbed there.

In the mechanism of inhibitor action so far considered, it has been assumed that monolayer adsorption is the position in which the organics reduce the rate of the cathodic or anodic reaction, and indeed in Figs. 12.42 and 12.43, two positions of octynol, lying down and standing up, are shown. Time-resolved automatic ellip-
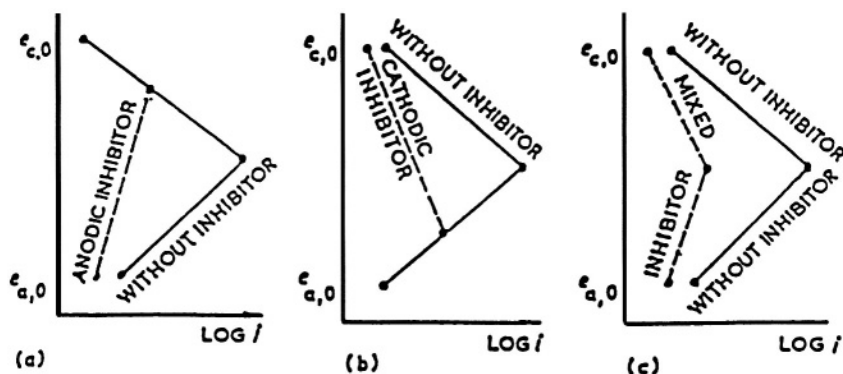
Fig. 12.49. Action of corrosion inhibitors. (a) Anodic inhibitors. Examples: Chromate, nitrite, molybdate, tungstate, orthophosphate, silicate, benzoate. (b) Cathodic inhibitors. Examples: $Ca(HCO_3)_2$, $ZnSO_4$, $Cr_2(SO_4)_3$, $NiSO_4$ phosphate, aminoethylene phosphate (AMP), Ag^{3+}, Sb^{3+} (on iron); Hg (on zinc). (c) Mixed inhibitors. Examples: organic inhibitors containing nitrogen and/or sulfur (e.g., amines, triazoles, thiazoles, alkylthioureas); inorganic inhibitors (e.g., arsenite, arsenate, selenate). (Reprinted from G. Ranglen, *Corrosion of Metals*, p. 165, 1985 with permission from Chapman & Hall.)

sometry has been used to examine what happens to octynol when it adsorbs on iron at temperatures (~75 °C) approaching that of boiling HCl. An important result is shown in Fig. 12.52. It is seen in the figure that several plateaus of the ellipsometric signal are observed as a function of time. The strength of these signals is proportional to the surface concentration of the inhibitor. The film is clearly growing in layers. Octynol assumes its upright position in the formation of what is a polymer film. Thus, the plateaus mount up to 5–6 layers of octynol, and this tightly bound polymer is evidently protective against the Cl^- ion and its tendency to enter the protective oxide on iron, thus causing its breakdown and a resulting increase in corrosion.

Such a result brings up the question of paint. Most exposed metal structures are painted to delay corrosion. Paint is the ultimate corrosion inhibitor.[8]

12.2.8. Nature of the Metal Surface in Corrosion Inhibition

In thinking about corrosion inhibition, it is implicitly assumed that one is dealing with a *metal*. This is fair enough in considering corrosion in acid solution, but the

---

[8]However, metals do corrode slowly, even when covered with paint, as every owner of a sufficiently old car knows. Pinholes in the paint start some of this (Fig. 12.25), but some paints also suffer the gradual corrosion of metals through them. Their water content is important.
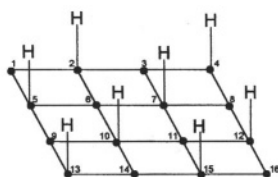
Surface plain with 8 adsorbed hydrogen atoms

Plain	Ave. Charge
1 ●	0.480
2 ●	0.123
3 ●	0.141
4 ●	0.155
5 ●	0.893
6 ●	0.050

Charge on atoms of the surface plain							
1	0.006	2	-0.950	3	0.006	4	-0.950
5	-0.950	6	0.009	7	-0.0950	8	
9	0.05	10	-0.950	11	0.023	12	-0.950
13	-0.95	14	0.029	15	-0.020	16	-0.950

Charge on each hydrogen atom: -0.09
Bond order Fe-H : 1.61

Fig. 12.50. Model of iron surface plane covered by adsorbed hydrogen. (Reprinted from P. Kutej, J. Vosta, J. Pancir, J. Macak, and N. Hackerman, "Electrochemical and Quantum Chemical Study of Dibenzylsulfoxide Adsorption on Iron" *J. Electrochem. Soc.* **142**(3): 833, 1995. Reproduced by permission of the Electrochemical Society, Inc.)

assumption tends to get less realistic as the pH increases. Indeed, in the real world, the corrosion situation is more likely to be concerned with the $6 < pH < 8$ conditions of natural waters, perhaps lowered to as little as pH 4 by acid rain.

A specific situation of great importance is that of the pipelines in the ocean beds under which lie large amounts of oil. In the North Sea situation, the oil arises from beneath the sea bed at temperatures that can range up to 60 °C, and an intentional CO_2 pressure is exerted on the oil in the pipelines, which brings the pH of the accompanying seawater layer between the oil and the pipe to a pH around 4. In this and similar circumstances, is it justified to continue to consider the corrosion of the metal of the pipeline as though the saline water was in contact with iron? Or should the surface to be considered be an oxide or even a carbonate?

Research to distinguish these alternatives (Kang, 1997) was carried out on the following basis. If the relevant surface is the metal, then it is known that the $\theta_{org}-V$ relation is parabolic around the pzc as in Fig. 12.53. However, if the relevant adsorption of the inhibitor is on a semiconducting oxide or carbonate covering the metal, little change in the coverage of the surface with organic inhibitors would be expected as a function of potential because in a semiconductor, most of the potential difference is
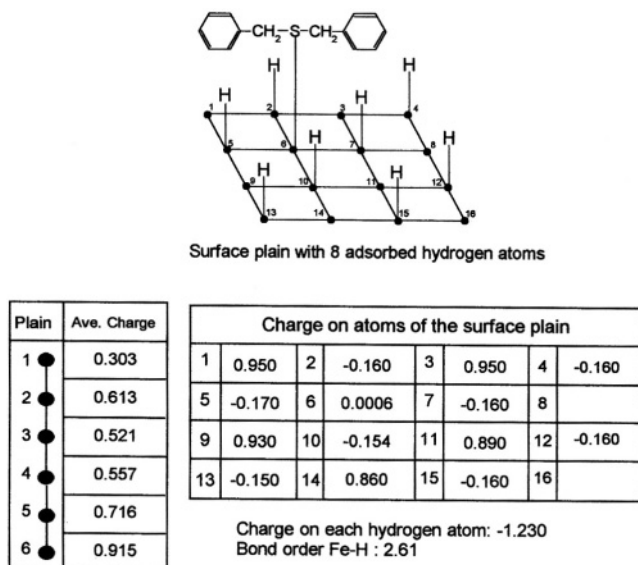
Surface plain with 8 adsorbed hydrogen atoms

Plain	Ave. Charge
1 ●	0.303
2 ●	0.613
3 ●	0.521
4 ●	0.557
5 ●	0.716
6 ●	0.915

Charge on atoms of the surface plain							
1	0.950	2	-0.160	3	0.950	4	-0.160
5	-0.170	6	0.0006	7	-0.160	8	
9	0.930	10	-0.154	11	0.890	12	-0.160
13	-0.150	14	0.860	15	-0.160	16	

Charge on each hydrogen atom: -1.230
Bond order Fe-H : 2.61

Fig. 12.51. Coadsorption of dibenzyl sulfoxide with hydrogen. (Reprinted from P. Kutej, J. Vosta, J. Pancir, J. Macak, and N. Hackerman, "Electrochemical and Quantum Chemical Study of Dibenzylsulfoxide Adsorption on Iron" *J. Electrochem. Soc.* **142**(3): 834, 1995. Reproduced by permission of The Electrochemical Society, Inc.)

within the solid, and the potential difference at the oxide/solution interface, that which would affect the adsorption of the inhibitor, would be small and ineffective.

Thus, the critical experiment to distinguish whether the inhibitor is adsorbing on the metal or on a protective layer is to find out whether the $\theta_{inhibitor}$–V graph is parabolic. The distinguishing experiment was done with ^{14}C labeled phenylalanine. Figure 12.54 shows the potential dependence of the adsorption of this compound on platinum (to show how the adsorption of this substance behaves on a metal) and on Fe in contact with a pH 3.8 phosphate buffer.

It is seen that the organic molecule adsorbs on the Fe in the same way with respect to potential as on Pt, there being a shift in the potential of the maximum corresponding to the difference of the pzc for Pt and Fe. Thus, it *is* correct to consider inhibitor adsorption in pH 4 as occurring *on* the metal of the pipeline rather than on an oxide or carbonate coating. It is possible that some oxide or carbonate exists on the metal surface in these situations of pH 4 (CO_2 saturated), but that the corrosion and its inhibition occurs through cracks in the film, which expose bare Fe to the solution.
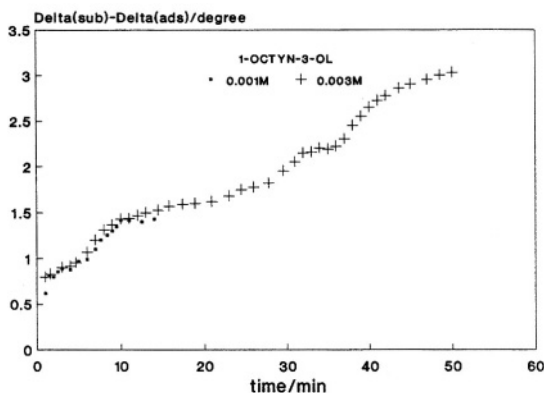
Fig. 12.52. Δ_{sub}-Δ_{ads} versus time relations for the adsorption of various concentrations of 1-octyne-3-ol on Fe at −0.42 vs. NHE at 73 °C. (Reprinted from J. O'M. Bockris and B. Yeng, *Electrochim. Acta* **36**: 1331, copyright 1991, Fig. 8, with permission from Elsevier Science.)

12.2.9. Green Inhibitors

Since the early 1990s, in the North Sea and in other areas of the world where platforms in the sea are used to extract oil, an environmental situation has arisen that challenges the use of the present corrosion inhibitors. When the oil from below the sea bed, with its seawater layer, is delivered onto the platforms, the accompanying sea water, containing much of the corrosion inhibitor, is ejected into the sea surrounding the platform. Since the majority of inhibitors in use at century's end are toxic to the surrounding sea life, the latter is assaulted. The long-term lethal effects of the inhibitors have led European governments to issue strictures on the amount of inhibitor that may be injected into the surrounding sea by each oil company. The limits are to be applied starting in the year 2000 and an aim of corrosion research during the 1990s has been to design organic substances that retain their efficacy as corrosion inhibitors but are several orders of magnitude less toxic than those now in use.

The basis of this work (Singh, 1996) has been to relate structures with diminished toxicity to their *lack* of affinity for the organic material (lipid structures) that is characteristic of the outer bodies of sea fauna. The actual toxicity of a compound depends how it interacts with critical enzymes within the living organism. To be able to affect such enzymes, the organic corrosion inhibitor must penetrate the lipid layer that constitutes the exterior of the organism. If structures

10⁻² M n-valeraldehyde
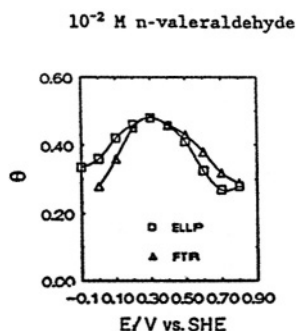
Fig. 12.53. Experimental coverage–potential relation for *n*-valeraldehyde on platinum as a function of electrode potential. Two things should be noted: First, adsorption–potential plots are roughly parabolic. Second, there is medium-good agreement between two entirely different methods, one based on the interpretation of ellipsometric data and the other on FTIR measurements. (Reprinted from J. O'M. Bockris and K. T. Jeng, *J. Electroanal-chem.* **330**: 553, copyright 1992, with permission from Elsevier Science.)

can be synthesized that have a low affinity for this lipid layer, their toxicity will be small.

Toxicity is measured by the concentration in mg liter^{-1} of a compound that causes the death of a certain percentage (usually 50 or 100%) of the test population of a chosen organism (e.g., silvery minnows) in a chosen time (e.g., 96 hours). For organic inhibitors, the higher the concentration needed to achieve a lethal dose of 50%, the less toxic the inhibitor. In Table 12.3 the actual lethal concentration (LC_{50}) (at 96 hr) is compared with that calculated by means of a quantitative structure–activity relation (QSAR). The basic calculation is that of the distribution coefficient of the inhibition of the primary alcohol octanol.
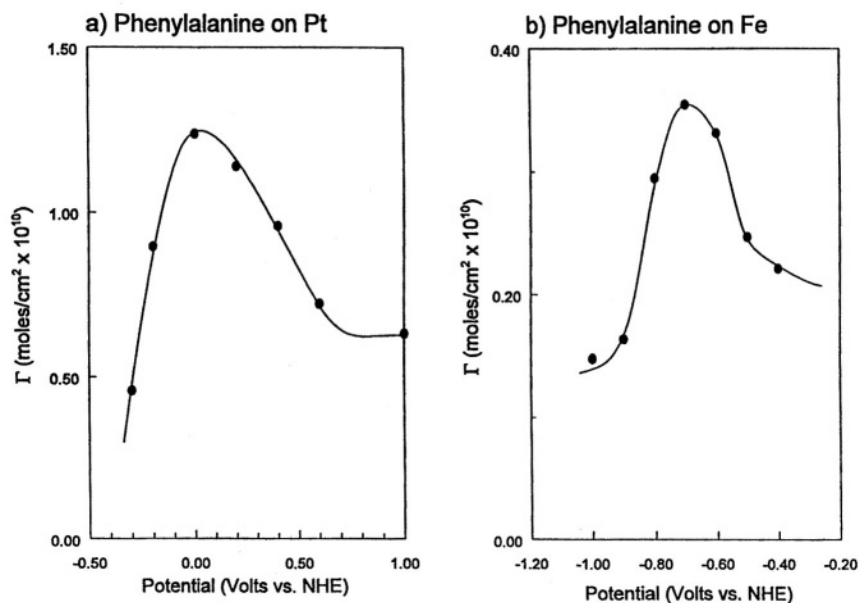
Fig. 12.54. Potential dependence of surface excess of phenylalanine at pH = 3.8 phosphate buffer. (From J. Kang, thesis, Texas A&M University, 1997.)

12.2.10. Looking Back on Some Methods by Which We Are Able to Inhibit Corrosion

The methods used to inhibit and effectively stop corrosion can be divided into two large groups depending on the corroding situation. If the body to be protected is in an infinitely large solution, typically the sea, then the methods have to rely upon what can be done electrochemically to the metal itself. Usually, as in the widely practiced cathodic protection, a circuit is fixed up in which the metal of the object to be protected is moved away from the corrosion potential in the cathodic direction (or electronation), thus reducing the anodic (or deelectronation) dissolution velocity, and hence the corrosion.

Alternatively, if the object to be protected consists of a container, such as a steam boiler or car radiator, then substances can be added to the solution inside the object and by their effect on the cathodic or anodic branch of the corroding reaction, slow it down so that the overall rate of corrosion is greatly reduced. Most of the substances that make up this latter class are organic compounds, often quite complex ones.

In designing such compounds, one has to think first, as to whether the corrosion rate is determined by the cathodic or anodic partial current of the two currents that make up the corroding couple. This is important because inhibitors are known which

TABLE 12.3
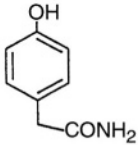**Green Corrosion Inhibitors: Ability to Predict Toxicity**
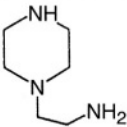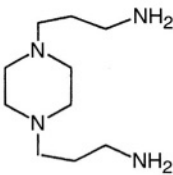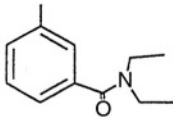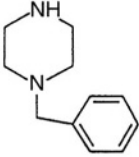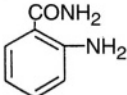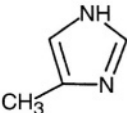
Structure	Exp. LC_{50} (96 hr) (mg/liter)	Estimated LC_{50} (96 hr) (mg/liter)
	2190	2804
	1480	5044
	3100	3319
	1130	733.5
	109.9	65
	48	122
	394	331

TABLE 12.3
Continued

Structure	ExpLC$_{50}$ (96 hr) (mg/liter)	Estimated LC$_{50}$ (96 hr) (mg/liter)
$C_{10}H_{21}NH_2$	1.01	0.227
	285	150

affect either the one type or the other. It is critical first of all to make sure which one is relevant.

Then there are rules about how much the organic adsorbs, and therefore inhibits as a function of potential. Organic substances tend to adsorb around the potential of zero charge (Section 6.9.3) of the metal concerned. However, aromatic organics particularly adsorb broadly around the pzc, up to 0.5 V in either direction, and so the relationship between the corrosion potential and whether it lies within the potential range of adsorption is easy to determine.

What is important about an organic and its potential strength as an inhibitor is the strength of binding of the organic to the surface. Aromatics like benzene derivatives, alicyclics like pyrroles, and sulfur-containing compounds are the best inhibitors, at least for iron-based alloys. The degree of inhibition also depends on concentration, and there is a wide range of concentrations at which a given inhibitor will completely cover the surface of the substance to be inhibited. Indeed, the concentration at which this occurs may go down as low as 10^{-6} M, or be as high as 10^{-3} M. The solubility of the organic plays an inverse role; the less the solubility, the more the tendency to adsorb.

Design engineering is at work and earlier an example was given of how, for dibenzyl sulfoxide, the strength of the inhibition can be estimated numerically. One way, a rather general one, of working out what structures will give what degree of inhibition uses the technique called a QSAR (see Section 12.10.9), a quantitative structure activity relationship.

Ab initio quantum mechanical calculations of binding of inhibitors to surfaces have been tried, but it is too much at the moment for the software to be able to give an answer in a reasonable time with such large molecules, and better to depend upon

quantum mechanical "indications" (for example, the HOMO–LUMO match of organic and metal) rather than an actual calculation of bonding.

In extreme cases, such as boiling HCl in contact with stainless steel, inhibitors can still be found that work very well indeed. Ellipsometric measurements show that they can form polymer layers five and six molecular layers thick—a kind of *in situ* formation of a paint film, built up by successive adsorption, layer upon layer.

It is often asked what the nature is of the surface in which the inhibitor adsorbs. It would be naive to assume that all actual steel surfaces have steel in contact with the solution. In some cases CO_2 is injected into the solution and this would tend to form carbonate films on the metal's surface unless the solution is more acid than, say, pH 4. Analysis of this problem shows that, indeed, inhibitors behave as though they are adsorbed upon the basic bare metal beneath any oxide or carbonate film. There is, however, evidence for films (which, wherever they exist would be protective), so that the actual inhibitor action is on bare metals formed in cracks and crevices between the zones of protective oxides and carbonates.

Finally, environmental restraints have caught up to corrosion inhibition science, too, and it has become necessary for situations in which the inhibitor (as with the oil platforms at sea) ends up by being ejected into the sea, to make it environmentally friendly so that minimal damage is done to the living inhabitants of the waters. Green inhibitors are those that have a poor affinity for octanol and settle more into water than into the organic substance. This makes them nontoxic (i.e., they would be less likely to penetrate into bioorganisms) but, of course, care has to be taken that they contain bonds (and sulfur bonds are very good at this for Fe) that compensate for their affinity for water and make them adsorb well onto the metal.

Further Reading

Seminal
1. Sir Humphrey Davy, *Phil. Trans. Roy Soc. London* **115**: 158 (1825). The first paper on cathodic protection (of British Navy ships).
2. J. O'M. Bockris and B. E. Conway, *J. Phys. Colloid. Chem.* **53**: 527 (1949). First established relation between hydrogen overpotential and corrosion inhibition.
3. E. L. Cook and N. Hackermann, *J. Phys. Colloid. Chem.* **55** :549 (1951). Adsorption as a prerequisite of inhibition.
4. A. C. MacKrides and N. Hackerman, *Ind. Eng. Chem.* **47**: 1773 (1955). How adsorption relates to inhibition.
5. I. N. Pictilova, S. A. Balezin, and V. P. Baranek, *Metallic Corrosion Inhibitors*, Pergamon Press, New York (1960). Details first patent for an inhibitor, issued to S. Baldwin, British Patent 2327 (advised use of molasses).
6. E. Blomgren, J. O'M. Bockris, and C. Jesch, *J. Phys. Chem.* **65**: 2000 (1961). The relation of the structure of the organic to adsorption and corrosion inhibition.
7. J. O'M. Bockris and P. K. Subramaniam, *Corros. Sci.* **10**: 435 (1970). The electrochemical basis for the stability of metals.

Review

1. N. Hackerman, *Langmuir* **3**: 922 (1981).

Modern

1. D. Rolle and J. W. Schultze, *J. Electroanal. Chem.* **229**: 141 (1987).
2. I. Macek, N. Hackerman, and Z. Haulas, *Proc. 7ᵗʰ European Symposium on Corrosion Inhibition*, p. 12 (1990).
3. J. O'M. Bockris and B. Yang, *J. Electrochem. Soc.* **138**: 8 (1991).
4. C. Vitozzi and G. D. Angellis, *Aquatic Toxicology* **19**: 167 (1991).
5. M. Chesallis, *Chemosphere* **22**:1175 (1991).
6. J. O'M. Bockris and K. T. Jeng, *J. Electroanal. Chem.* **330**: 541 (1992).
7. S. N. Raicheva, B. V. Aleksiev, and F. I. Soholava, *Corros. Sci.* **34**: 343 (1993).
8. H. S. Rosenkranz, E. J. Matthews, and G. Klopman, *Ecotoxicology Env. Safety* **25**: 296 (1993).
9. P. Kutej, J. Vosta, J. Macak, and N. Hackerman, *J. Electrochem. Soc.* **142**: 829 (1995).
10. P. Kutej, J. Vosta, J. Pancir, and N. Hackerman, *J. Electrochem. Soc.* **142**: 1847 (1995).
11. B. Yang, H. Zheng, and D. A. Johnson, *The Inhibition of H Permeation in Corrosion*, Paper 271, National Association of Corrosion Engineers, Houston, TX (1997).
12. P. Mutumbo and N. Hackerman, *J. Solid State Electrochem.* **1**: 194 (1997).

12.3. THE PROTECTION OF ALUMINUM BY TRANSITION METAL ADDITIONS

12.3.1. Introduction

One of the methods of protecting metals from corrosion mentioned in the treatment of inhibition is anodic protection. Discussion of this topic was delayed until now because it leads to a subject—passivity—that merits a section on its own. But before we get to that, there is a special case to deal with, that in which a few percent of transition metals, added to aluminum, provide it with a surprisingly high degree of protection. This subject is presented here because although it relates to inhibition, protection against the ravages of corrosion, it is close to passivation (Section 12.4) and is quite different in mechanism from that by which certain organic molecules inhibit corrosion.

Aluminum must rate as the second most technologically important metal. Its ores (bauxite at present and clay when that is exhausted) are abundant; it has about half the specific gravity of iron, and its metallurgical properties are good enough to be the principal component of the alloys used in the bodies of aircraft and their much-stressed wings. Aluminum corrodes easily at extremes of pH but has a substantial pH range (4–12) in which protective oxide films exist, a very important fact when surfaces are brought into contact with sea water or salt spray. These oxide films, which protect aluminum in the middle pH ranges are, however, subject to attack by Cl^-.

An important advance (Davis and Moshier, 1990) occurred when it was found that quite small quantities (5 mol % and less) of transition metals (e.g., Ta, Mo, and W) conferred a surprisingly high degree of protection on Al. Thus, in rough terms, the addition of small amounts of these materials to Al cut the rate of corrosion of Al by around 100 times, and the time to breakdown under constant electric field across the protective oxide layer is increased by about 10 times.

The question is: Why? A naive first model of the effect of adding a novel component to a metal less likely to dissolve than the host metal would suggest a slowdown of corrosion proportional in rate to the fraction of "blocking component present" in the surface (i.e., 10% of a nonactive component would cause a 10% reduction in corrosion rate). That Ta, Mo, and W cause a reduction in corrosion rate of 100 times when present to an extent of less than 10% hints at a special mechanism.

A compact way of illustrating these phenomena is to exhibit the anodic polarization curves for Al and four of its alloys. As the metals are added, the potential needed to cause breakdown and the subsequent pitting is shifted up to as much as 0.85 V anodic to the potential of breakdown in pure Al (Fig. 12.55).

12.3.2. Some Facts Relevant to the Transition Metal Effect on Inhibiting Al Corrosion

The effect stated above is observed in Cl^- *containing solutions.* It is well known that Cl^- ions are particularly strong attackers of protective films (Mizuno, 1986).
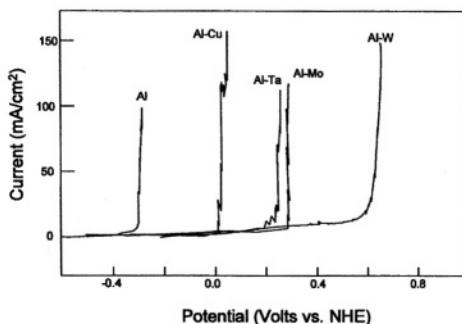
Fig. 12.55. Polarization curve for Al and Al alloys in pH 8.4 borate buffer. $Cl^- = 0.01$ *M.* Sweep rate = 1 mV s^{-1}. (Reprinted from J. O'M. Bockris and J. Y. Kang, "The Protectivity of Aluminum and its Alloys with Transition Metals," *J. Solid State Electrochem.* **1**: 19, 1997, with permission from Springer-Verlag.)

Hence, the presence of Cl^- on the surface of Al or Al alloys would be at least a precursor to breakdown. A very telling experiment (Kang, 1997) shows that the potential for the adsorption of Cl^- on the Al to which the transition metals have been added is shifted substantially in the anodic direction (Fig. 12.56).
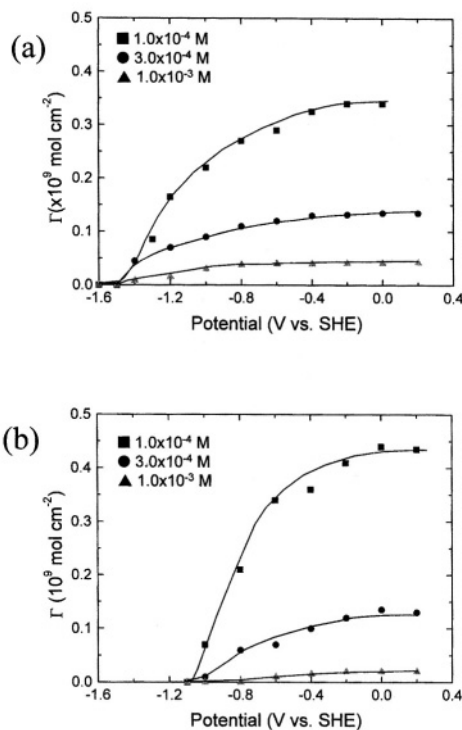

Fig. 12.56. (a) Adsorption of the chloride ion on Al as a function of applied potential in pH 9.2 borate buffer containing different chloride concentrations. $Cl^- = 1 \times 10^{-3}$ M ($\triangle$), 3×10 M ($\square$), and 1×10^{-4} M (O). (b) Adsorption of the chloride ion on an Al-Ta surface as a function of applied potential in pH 9.2 borate buffer containing different chloride concentrations. $[Cl^-] = 1 \times 10^{-3}$ M ($\triangle$), 3×10^{-4} M ($\square$), and 1×10^{-4} M (O). (Reprinted from J. O'M. Bockris and J. Kang, "The Protectivity of Aluminum and its Alloys with Transition Metals," *J. Solid State Electrochem.* 1: 21, 1997, with permission from Springer-Verlag.)

Hence, for whatever reason, the effect of alloying is to delay the potential at which the surface charge becomes positive (hence enable the Al surface to adsorb Cl^- anions) until a much more positive one (about 0.4 V more positive if one compares the potential at which adsorption begins for Al-Ta with that at which it begins in Fig. 12.56.

This result gave rise to a desire to measure the potential of zero charge for Al and its alloys[9] with transition metals. This quality can be determined fairly easily for metals in contact with solutions (Trasatti, 1990), but in the present case it is protective oxides (i.e., semiconductors) in contact with the solution, and the determination of the corresponding potential of zero charge is much more difficult. It is possible, however, by the use of impedance measurements to determine the flatband potential (Section 6.10.1), and this potential would be identical with the potential of zero charge were it not for the specific adsorption of OH^- and Cl^- ions, which occurs on the surface of these oxide-covered Al alloys at certain potentials and concentrations. However, it is possible (Kang, 1997) to correct for the anionic charge adsorbed at the interface and thus to go from a measurement of the flatband potential obtained by plotting $1/C^2$ against V—the Mott–Schottky plot (Sec. 10.5) to the potential of zero charge (the uncertainty in the calculation is ± 1 μC cm^{-2}).

An important correlation between the pitting potential and the pzc results and is shown in Fig. 12.57. It is seen that the pitting potential (an arbiter of the ease of corrosion) is directly related to the potential of zero charge of the alloy electrodes with their oxide-covered surfaces. The considerable shift of pzc ($\sim$0.8 V) with the alloys correlates excellently with the shift of potential at which the current through the film curves up sharply, i.e., at which the film breakdown has led to pitting and thus easier transport of ions (and charge) through the film.

Thus, the pzc of the various alloys—and the resulting shift of potential at which the film-breaking anion Cl^- can adsorb—can be correlated. Breakdown of the protecting oxide layers on Al and its alloys does not occur until Cl^- adsorbs. Hence, the protectivity bestowed by the alloys would be interpreted as being due to the fact that alloy formation shifts the pzc in the anodic direction and prevents Cl^- adsorption (with its destructive effects on the oxide film) until more and more anodic potentials, are reached greatly slowing down the Cl^- dependent corrosion. In practice, the alloys are at potentials negative to the potential of Cl^- adsorption (and hence film breakdown and resulting corrosion). With pure Al, the fact that the pzc occurs at relatively negative potentials means that in many practical situations, the surface of the Al will be positive to the potential of its pzc and hence Cl^- adsorbing and oxide destroying.

However, before this solution to the problem of the greatly increased protectivity of transition metal alloys of Al is accepted, it is relevant to ask about the structure of the oxide film. Are the alloying components preferentially adsorbed at the interface,

---

[9] In this section the Al systems containing 5–10 mol% percent of certain metals are called alloys. However, even at this small concentration, the systems are supersaturated. In the normal temperature range and at practical times, they nevertheless act as stable alloys.
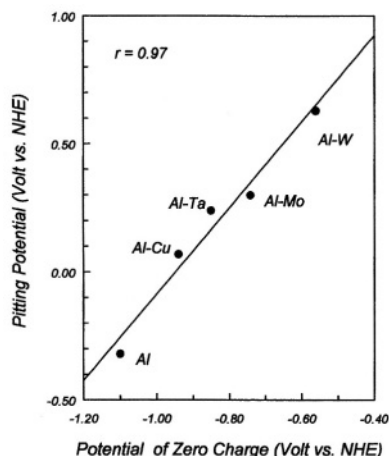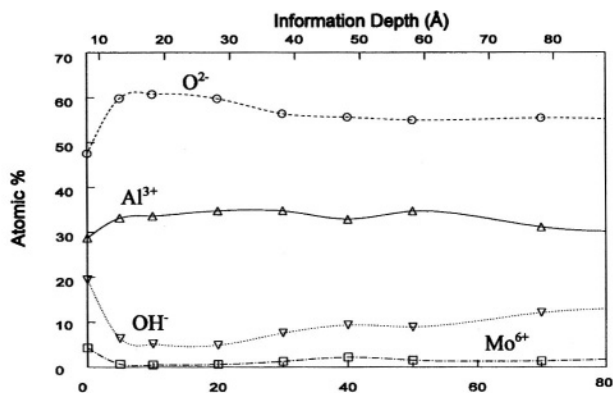
Fig. 12.57. Correlation between pitting potential and pzc. (Reprinted from J. O'M. Bockris and J. Kang, "The Protectivity of Aluminum and its Alloys with Transition Metals," *J. Solid State Electrochem.* **1**: 28, 1997, with permission from Springer-Verlag.)

thus giving rise to the strong shift of pzc established? The oxide's structure can be analyzed by using two methods, X-ray photoelectron spectroscopy (XPS) (Section 6.2.5.2.b) and ion scattering spectroscopy (ISS). By using a beam of argon ions to sputter away the oxide layer by layer and knowing the average depth registered by XPS (about 20 Å), it is possible to analyze the materials present at all depths of the oxide films. Correspondingly, the ISS method gives knowledge of the first layer of the oxide exposed by the sputtering and thus also allows one to know the composition at any depth. In Fig. 12.58 the XPS analysis of O^{2-} Al^{3+}, OH^- and the allowing element is seen as a function of depth. In Fig. 12.59 the OH^- concentration is plotted as a function of depth. Two marked results come out of these analyses.

    1. The alloying element, which causes so much change in the corrosion behavior of the alloys, is scarcely enhanced on the surface. In fact, for Ta and Cu, the surface concentration is somewhat decreased from that in the bulk. The maximum surface concentration of the alloying transition metals is 5 mol % (one atom in 20 on the surface).

    2. The OH^- concentration (earlier thought to be restricted to a thin-surface layer) does in fact extend throughout the whole oxide layer (total thickness ~ 60 Å). This fact will prove mechanistically important, as will be seen.

(a)

(b)
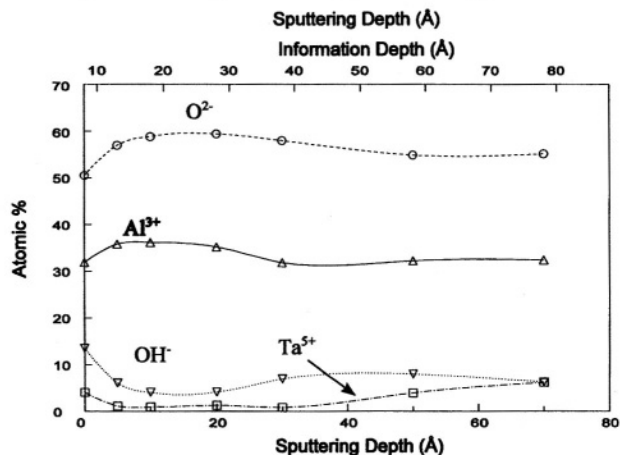
Fig. 12.58. XPS atomic concentration profile of the passive film of Al alloys as a function of sputtering depth: (a) Al-Mo, (b) Al-Ta, (c) Al-W, (d) Al-Cu. Passive film formed in borate buffer at +1.0 V (vs. NHE). (Reprinted from J. O'M. Bockris and J. Kang, "The Protectivity of Aluminum and its Alloys with Transition Metals," *J. Solid State Electrochem.* **1**: 23, 1997, with permission from Springer-Verlag.)

Another way of showing the structure of the oxide film in the presence of alloys is a 3D presentation (L. Minevski, 1993). Figure 12.60 clearly shows the two types of O, i.e., O of Al_2O_3 and O of OH^- (largely in $Al(OH)_3$) which itself lies near the surface.[10]

---

[10]This suggests that the rest of the OH that decreases in concentration as one leaves the interface continues to be held in some other structure. L. Minevski (1993) suggested that this was in a "fibril," a chain of AlO-OH that has been found in some structures to exist in rodlike form within oxides (Alwitt, 1976).

(c)
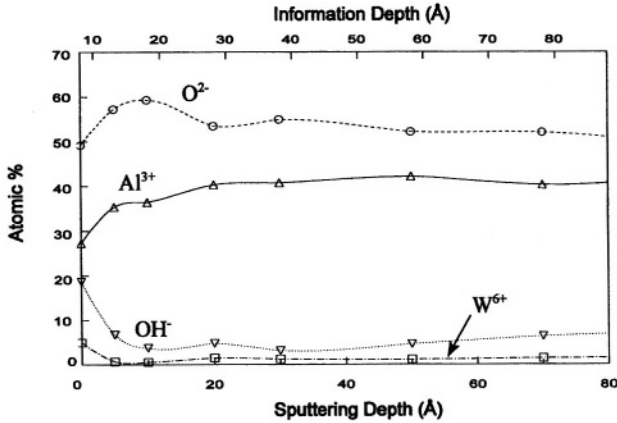
(d)
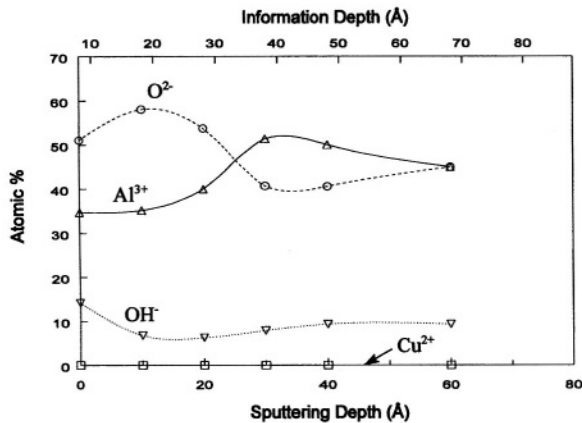
Fig. 12.58. *Continued*

12.3.3. The Model by Which Tiny Concentrations of Transition Metal Ions Retard Corrosion of Al

One model has already come out of the experimental results. The shift of the pzc in the anodic direction as the alloying metals are added shifts Cl^- adsorption in the anodic direction and leaves more and more of the practical potential range in which no Cl^- can adsorb, because the surface carries a repelling negative charge. However, there are two clues that suggest that Cl^- adsorption (and its consequent attack on the integrity of the protecting film) is a necessary part of the inhibition but that there is some other factor that delays Cl^- entry even after it has adsorbed. The clues are related to the structure of the layer (Fig. 12.61).

(a)

(b)
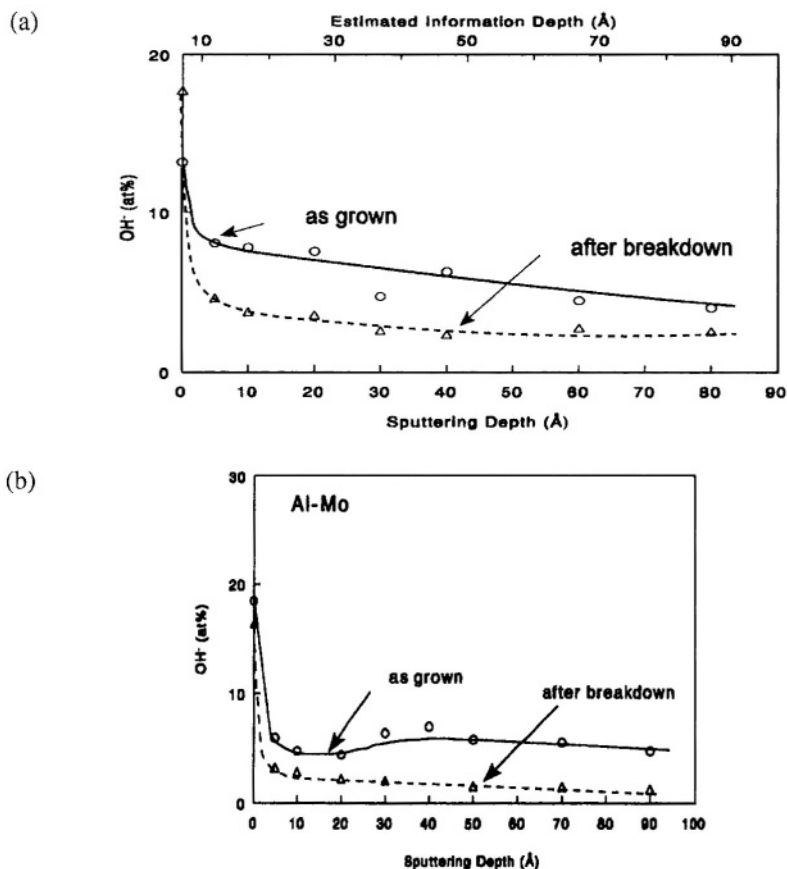
Fig. 12.59. Depth profile of the concentration of OH⁻ for Al and Al alloys: (a) Al, (b) Al-Mo, (c) Al-W, (d) Al-Ta. (Reprinted from J. O'M. Bockris and J. Kang, "The Protectivity of Aluminum and its Alloys with Transition Metals," *J. Solid State Electrochem.* **1**: 24, 1997, with permission from Springer-Verlag.)

1. The OH⁻ concentration is not only spread throughout the oxide (Alwitt–Minevski fibrils) but decreases substantially on breakdown. Thus, the absorption of Cl⁻ ions (for their absorption must follow their adsorption to disrupt the oxide) must have some connection with the OH⁻ in the oxide.

2. The rate of the passive film breakdown at Al-Ta is one-tenth that of Al at the same field strength within the oxide. But that is not what would be expected. The concentration of the alloying element is too small (1 place in 20) to retard the

(c)
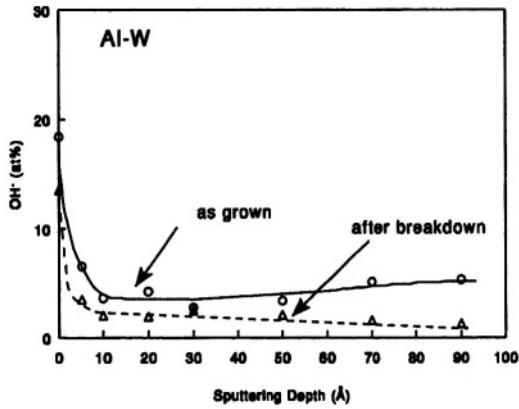
(d)
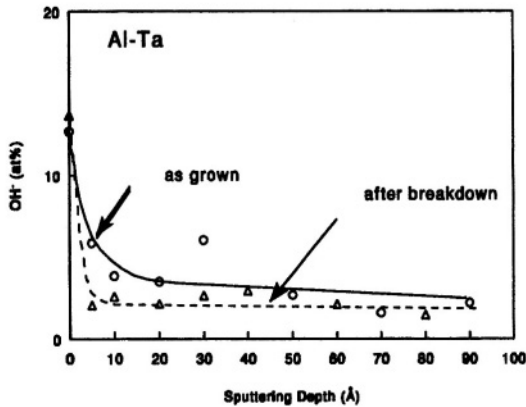
Fig. 12.59. Continued

general rate of Cl⁻ entry and passage through the film toward its interface with the metal if that entry took place everywhere on the film surface after adsorption had occurred.

Thus, a satisfactory model must involve functions in addition to that of the influence of the change in pzc with alloying on Cl⁻ adsorption on the surface of the protecting oxide. Minevski (1993) had suggested that the Cl⁻ needed specific "entry points" into the oxide and these would be provided where the fibrils lying throughout the film met the surface. Hence, if the diffusion path of Cl⁻ involved a place exchange with OH into fibrils, the function that the OH groups play would be explained.

This view can be expanded. There can be other origins of flaws in the oxide surface that could be entry points (Wood and Richardson, 1973). The alloying cations can diffuse into the double layer outside the oxide on the solution side of the interface with the oxide and readsorb as anions when the surface becomes positively charged. They
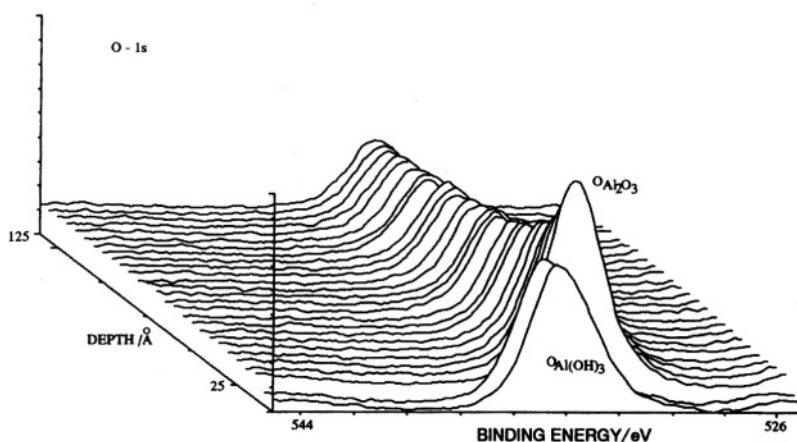
Fig. 12.60. 3D-XPS depth profile of passive film formed on Al for pH 8.4, at $V = 0.4$ V for 1 hr distribution of two different O species (as indicated) *before* the breakdown occurred. (Reprinted from J. O'M. Bockris and L. Minevski, *J. Electroanal. Chem.* **349**: 388, copyright 1993, with permission from Elsevier Science.)

would be most likely to deposit at crystal imperfections. Thus, a critical function of the transition metal elements is evidently to *block* entry sites and retard Cl^- entry, even after it has adsorbed, and therefore stop the depassivation and film breakdown. This would occur only after the adsorbed Cl^- has surface diffused so the blocked entry points have displaced the blocking entities and then begin diffusion through the oxide, exchanging with the OH^- ions therein.
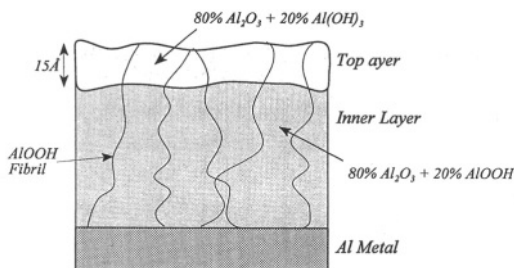
Fig. 12.61. The structure of the passive film of aluminum. (Reprinted from J. O'M. Bockris and J. Y. Kang, "The Protectivity of Aluminum and its Alloys with Transition Metals," *J. Solid State Electrochem.* **1**: 30, 1997, with permission from Springer-Verlag.)

To summarize the mechanism of the effect of transition metal ion alloys in inhibiting the corrosion of Al alloys:

1. The presence of the alloying elements causes a shift of the pzc in the anodic direction. Adsorption of the aggressive anion, Cl^-, from the solution onto the oxide (prior to its penetration and destruction of the film) cannot occur until the potential of the Al alloy is positive to the (anodically shifted) pzc in the alloys. But in many practical situations, the alloy surface remains negative to the positively shifted pzc, so that Cl^- does not adsorb.

2. This is a necessary but insufficient model to explain the facts which include, e.g., the fact that the *rate* of passive film breakdown is so much less on Al-Ta than on Al, although Cl^- is adsorbed on each (and the differing internal field strength in the oxide has been accounted for). Hence, another surface factor that is dependant on the transition metal ion must control the relative velocities of breakdown due to Cl^- on the various alloys. The Cl^- has to have entry points in the protecting surface oxides (specific models of these are suggested) and it is the removal of these blocking mechanisms that controls the rate of entry of the destructive Cl^- into the protecting oxide.

12.4. PASSIVATION

12.4.1. Introduction

Rust has been recognized for millennia and films of oxides on metals since the 1800s. However, for about 150 years, a tantalizing phenomenon remained unexplained. Some metals, suitably treated, enter a state expressively known as "passive." Their normal reactivity to outside stimuli (contact with dissolving acids, for example) is quelled. A reason for this was hard to find for more than 150 years. None of the methods generally available before about 1950 detected anything on the metal, and yet its character had changed dramatically.

A breakthrough came with the work of a Norwegian scientist, Tronstad,[11] who in 1933 for the first time, applied to passive film a method discussed by Raleigh in the nineteenth century—ellipsometry. This method works because it changes the properties of polarized light and does not depend on interference phenomena. Thus, it is not limited in the thickness of the film it can detect by the wavelength of light (about 5000 Å for visible light). Because of its mode of action (changing the polarization properties of light), it can be sensitive even to films of monolayer thickness. Tronstad did not go so far as to look at a passivating metal in solution under potential control (Reddy,

---

[11]Tronstad, who worked on passivity and ellipsometry with the British scientist Winterbottom, lived in the remote northerly city of Trondheim, just below the Arctic Circle. When the Nazis invaded Norway in 1940, he volunteered to be part of the Norwegian ski troops who resisted the German troops in the mountains surrounding his town. He did not return from his first encounter with the invaders.

1964), but removed the metal that had undergone the mysterious change to the passive state from solution and looked at its surface after the solution had been washed off. He found the secret. A protecting film was indeed there. The reason it had gone unexamined for so long was its thinness; it was only 30–40 Å thick.[12] But this degree of oxide gave apparently enough protection to entirely change the properties of the metal on which it sat. Thus, formation of a "passive layer" is in fact a mode of corrosion protection. Because passive films sometimes get formed spontaneously between metallic surfaces and their environment, it is not too fanciful to call passivity nature's method of corrosion inhibition.

We should reflect a bit about the larger meaning of passivity for the civilization built up since the seventeenth century. Only noble metals are stable in moist (i.e., normal) air. All the rest of the metals are not in their stable equilibrium state when exposed to oxygen in the presence of water. For electrochemical reasons that readers of this book now well understand (Section 12.17), they tend to corrode away. Why is it, then, that we can construct buildings, ships, and planes that last for something like 10, 5, and 4 decades, respectively, before they succumb to the electrochemical action of the surrounding moist air? It is clearly because of protective (mostly oxide) films. Perhaps some of these films are not the mysterious *passive* films, but they are oxide films and they do protect. So, as we explore what is known about the best of them—the passive film—it is worth recalling that we are investigating a vital link in the factors

---

[12]This was not the view long espoused by H. H. Uhlig, who was in the metallurgy department at MIT from 1936 until 1972. During this period, he was the foremost academic in the development of corrosion science in the United States and had a long-standing collegial relation with his opposite number at Cambridge University in England, Ulrich Evans. Much of the academic development of corrosion science in the United States passed across his desk, and many of the technological problems. Uhlig worked for many years on passivity and was the author of the idea that the critical film giving rise to the passivated state is a monolayer of oxide, a view that lost its hold on the corrosion community because the ellipsometric work during the 1960s found measurements in the tens of angstroms for iron and its alloys. Nevertheless, Uhlig maintained, with some justification, that the passivation potential itself was associated with a monolayer (the 30–40 Å was after growth).

Herb Uhlig was the advisor and intellectual father of many U.S. scientists who are at work in numerous companies in the United States. Among the best known of his students were Milton Stern, originator of the Stern–Geary method for the measurement of corrosion, and Winston Revie, who made the first Auger measurements on passive films.

Uhlig was thermodynamically oriented and for him corrosion occurred because of the difference in the "Galvanic potentials" of the components making up a corroding surface. He believed that corrosion originated from the presence of inclusions, detected or not, in the corroding material.

When Carl Wagner came to the metallurgy department at MIT in the early 1950s, an awkward nonequilibrium problem existed between the dour, frowning German kineticist and the sociable, smiling American thermodynamicist. The problem was never resolved although their offices in the same department were a few moments apart.

Uhlig led an active extracurricular life, both as a family man and hobbyist. His delight was in meteors, which he collected. During his life he made notable contributions to knowledge in this area, although he is mainly remembered for his contributions to passivity.

that make possible our technological civilization, which is made up of machines containing metal parts.

12.4.2. Some Definitions

In spite of the widespread practical importance of the protection of metals by their oxide layers, the strict scientific definitions connected with them have to involve laboratory experiments. The appropriate relation is shown in Fig. 12.62. Thus, if a suitable metal (e.g., Fe) is placed in the grip of a potentiostat[13] (see Fig. 12.63) and the potential is gradually made more anodic, after a period of exponential increase there occurs a maximum in the current–potential curve, after which (on further changing the potential) the current, instead of continuing the exponential increase predicted by the Butler–Volmer equation, unexpectedly collapses to a low value (Fig. 12.62). The potential at which that fall occurs is called the *passivation potential.* The metal is on the way to becoming passive, but the true state of passivity is not reached until the current density has stopped falling and remained at a kind of "trickle charge" value (about 2–3 $\mu A\ cm^{-2}$ at 25 °C for Fe in pH 8, for example), remaining constant at this low value even though the electrode potential is moved substantially in the anodic direction.

Because the association of the expression "passivation potential" with the potential of the maximum leaves something to be desired as a definition of passivation, another potential, the Flade potential, is sometimes used to define the phenomenon of passivation. This potential can be described quite loosely (see Fig. 12.62) as the potential at which (while moving the potential back in the cathodic direction) the current begins to increase from the low trickle charge value.

The precise definition of the Flade potential needs a little more description. Thus, if an electrode containing a passive layer is allowed to float free (the potentiostat is disconnected) and its potential is measured and plotted against time, it will fall (become less positive) and then attain a plateau on the potential–time graph (Fig. 12.64).

After a time (a few minutes in 0.5 $M\ H_2SO_4$ for Fe), the potential undergoes a sharp decline in the cathodic direction, and it is the potential at the beginning of this decline that is called the *Flade potential* (see Fig. 12.64). It is accepted that the protecting passive layer undergoes reduction when allowed to float free of potential control and that the sharp change in potential signifies the removal of the film and contact between the bare metal and the solution. Thus, the Flade potential is a kind of depassivating potential.

12.4.3. The Nature of the Passive Layer

Although the existence of very thin passive layers was first established by means of ellipsometry (Tronstad, 1933; Reddy, 1964), until 1973 there was no understanding

---

[13]The rate at which the potentiostat moves is very important in passivity measurements. The films take time to form and grow. The appropriate sweep rate will be slow, ~1 mV s^{-1}.
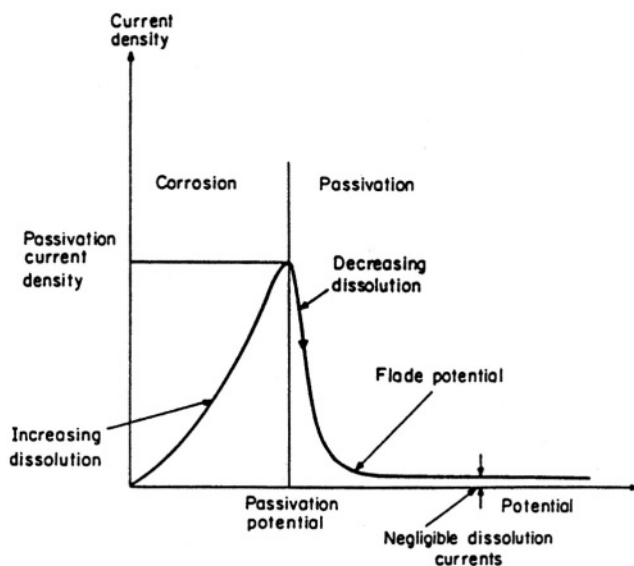
Fig. 12.62. As the potential of steel is increased, the current initially increases, reaches a maximum value, and starts sharply decreasing to negligible values.
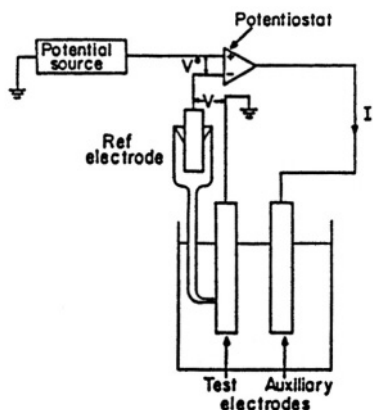

Fig. 12.63. In a potentiostatic circuit, an imposed potential difference V^* drives the current from the auxiliary electrode to the test electrode until the potential difference V becomes equal to V^*, so that the total input into the potentiostat falls to zero.
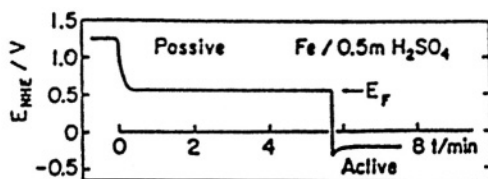
Fig. 12.64. Spontaneous decay of passivity as a function of time, showing the Flade potential of iron in 0.5 M sulfuric acid.

of the essence of what makes a passive film so protective. The breakthrough came in the interpretation of the results of Mössbauer spectra examination by O'Grady and Bockris (1973).

Mössbauer spectra concern the emission and resonant absorption of nuclear γ-ray emission given off by certain radioative isotopes, e.g., ^{57}Fe. The source of the rays is mounted on a platform that moves to and fro (0–2 mm s^{-1}), presenting to the specimen incident radiation over a narrow band of energies (the basic energy of the radiation, very slightly changed by 1/2 mV^2). At a certain very sharply defined energy, the radiation is absorbed by the passive layer on the electrode. The specific energy of absorption of the γ emission gives critical, identifying information on the chemical character of the oxide present—that constituting the passive film. Information on the properties of the passive layer can be obtained because the nuclei of the Fe atoms (hence the frequency of the resonant absorption) are affected by the character of their surroundings. The point of making these spectroscopic measurements specifically by Mössbauer spectra is that the changes in the nuclear γ emissions caused by changes in the surrounding force field (e.g., as Fe in FeO or differently, Fe in Fe_2O_3, etc.) are very small indeed and it is only the Mössbauer spectra that have the sensitivity (i.e., the tiny line width) to be able to detect and even to quantitatively measure the spectral changes which by their character identify the oxide that makes up the layer.

The normal ways in which Mössbauer lines are interpreted (in searching for the identity structure of the material under examination) is to compare the frequencies (actually the rate of motion of a carrier containing the source of γ emission, e.g., ^{57}Fe) at which resonance occurs[14] with the corresponding Mössbauer spectra of various

---

[14]The γ-ray emission from ^{57}Fe has a fixed frequency. However, the specimen emitting this γ-radiation characteristic of ^{57}Fe is being moved, repetitively, from 0 to 2 mm s^{-1}, and the kinetic energy of this movement has to be added to the energy of the γ-emission from ^{57}Fe Now the extra energy of $\frac{1}{2}$ mV^2 passes through a range of energies that are added to the basic ^{57}Fe energy. Hence the possibility arises of covering the change in the absorption frequency of the Fe atom caused by the passivating oxide, whatever its character may be. The difference between the absorption frequency of the Fe covered with the passive oxide and that of the original ^{57}Fe (the isomer shift) identifies the nature of the oxide that caused it.

possible alternative substances to see where there is a match. The change in frequency between that of the resonance absorption in the passivated specimen with its Fe nuclei altered by their new surroundings and that of the unaffected pure ^{57}Fe nucleus is called *the isomer shift.*

When the Mössbauer method was first used to examine the structure of the passive film, the values of this isomer shift were puzzling, for they did not fit the characteristic frequency shifts caused by any of the iron oxides, including that of γ Fe$_2$O, which had been earlier suggested as the identity of the passive layer oxide. O'Grady[15] (1971) observed that the values in the passive film were strikingly similar to those that Prados and Goode had found in work on iron-containing polymers. This was the clue that gave rise to the suggestion [verified by the EXAFS measurements of Kruger and Reveiz (1978)] that the essential characteristic of the passive film—why it does not dissolve in acid like others of similar stoichiometry—is its amorphous, chainlike nature. Hazony (1968) had published a plot of the dependence of isomer shifts on the hydration number, and the application of this relation to the results of their Mössbauer spectra enabled O'Grady and Bockris to determine the water per Fe atom (one water per Fe atom). The first model for the hydrated structure of the iron oxide polymer that makes up the passive layer of this best-known passive film is shown in Fig. 12.65.

The work of O'Grady and Bockris (1973) in discovering that the key characteristic of passive layers originates in their amorphousness required confirmation by an alternative method. This soon came when Revie, Baker, and Bockris (1974) applied an ultrahigh voltage (uhv) technique for the first time in electrochemistry to measure the difference between the Auger spectra of a passivated layer of Fe and that of one that had been broken down in the sense that it had been depassivated by being exposed to a solution containing Cl$^-$ ions. From such measurements they discovered a telling fact. When the iron was in the passivated state, the O/Fe ratio was 2:1 (=Fe$_2$O$_3$H$_2$O), but after depassivation, the ratio had changed to 1.6 to 1. Thus, if the essential point about the structure of the passive layer is amorphousness *caused by hydration* (Fig. 12.65), then of course the depassivated film has lost its water. But Fe$_2$O$_3$ is indeed equivalent to an O/Fe ratio of 1:5, i.e., close to the Auger result of 1.6.

Ellipsometry has already been cited as the method that first established that passivity was associated with the presence of a very thin oxide film. Later it was measured (Jovancicevic, 1987) by means of ellipsometric spectroscopy and it was thus found possible to evolve an extinction coefficient that varied characteristically with the wavelength at which the ellipsometry was carried out (Fig. 12.66). The interpre-

---

[15]Discovery and the euphoria that goes with it in successful research is well illustrated by the actions of Dr. William O'Grady, at the time a graduate student at the University of Pennsylvania. He entered his advisor's room precipitously without no knocking and exclaimed "ITS AMORPHOUS" in a loud and ringing voice. He had returned from the library after discovering the work of Prados and Goode on the isomer shifts for amorphous Fe-containing polymers. Research scientists know such moments; they occur rarely enough and are usually the result of many months (and even years) of very hard work in which nothing seems to be happening.
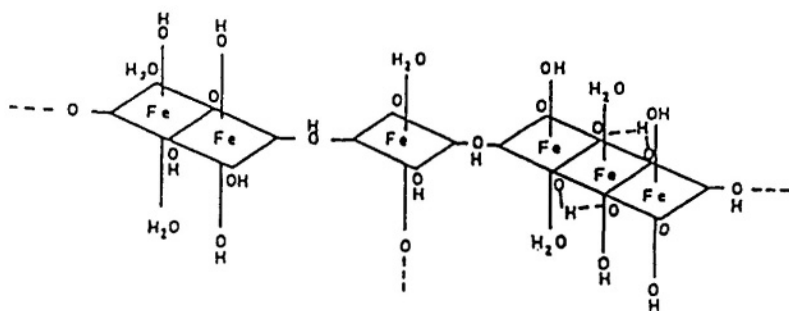
Fig. 12.65. A polymeric film of hydrate iron oxide consistent with Mossbauer spectra of passive film on iron.
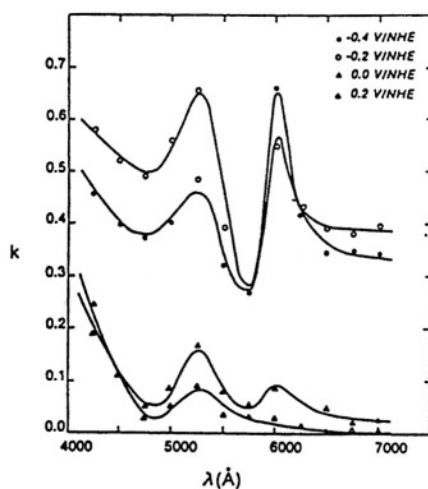

Fig. 12.66. Extinction coefficient spectra of iron for different potentials. (Reprinted with permission from V. Jovancicevic, R. C. Kaintla, Z. Tang, B. Yang, and J. O'M. Bockris, *Langmuir* 3: 388, copyright 1987 American Chemical Society.)

tation of these ellipsometric spectra indicated a degree of hydration of 80% of one water molecule for Fe_2O_3, which in turn supports the essential role that hydration plays in causing the passive layer to be amorphous and thus protecting.[16]

12.4.4. Structure of the Passive Film

It is the presence of H and OH groups that provides the essentials of the *passive* film and the reason that it is amorphous. The amorphous character leads to non-stoichiometry. There is a deficiency in protons. Passive layers contain a corresponding gradient of ionic species (Cahan and Chen, 1992).

Much of our knowledge of passivity is concerned with the passivity of iron. The reason of course is that this is the most important layer technologically. However, there is an increasing amount of work being done on other passive layers (for Al, see Section 12.3) and on Cr and Ti. For details of the structure on Cr and T, see the further reading sections.

12.4.5. Depassivation

Studies of the structure of passive layers are eventually of technological value only if they can substantially delay the breakdown of that passive layer which is so important to the stability of the metal it protects. As far as the all-important iron and its alloys are concerned, the polymeric oxide model, with the part played by water in putting together the polymer elements, seems to be the most consistent with the facts. In considering its breakdown, one generally discusses this in terms of the effects of Cl^- adsorption, but there are other ions (I^-, Br^-, SO_4^-) that also cause depassivation.

As traced out in detail for the breakdown of the protective (not passive) layers on Al (Section 12.3), the adsorption of the anion is the first and most vital step. It seems, at least for the Al alloys with transition metals, that the entry of the depassivating anion into the film is held up by some surface process between the adsorbed ion and a surface defect, but at any rate the anion does eventually go in and diffuse through the film. In this diffusion, it meets the OH species in the film and displaces them (a fact proven by several methods). In the Fe system, it forms a Cl-containing Fe complex that leaves the system and can be detected outside the depassivating system. A model of Cl^- going inside the oxide as an important step in depassivation was first suggested by Pryor (1965), but the proof was not given until the radiotracer work of Mizuno (1986).

Figure 12.67 shows the dependence on time of Cl^- diffusion into the passive layer on Fe. Thus, when a passive layer of Fe in pH 8.4 borate buffer is exposed to

---

[16] Why is it assumed that an *amorphous* film will be protecting, i.e., passive? It is because depassivation—the breakdown of the film—is associated with the easy passage (by means of electrodiffusion) of Fe from the metal through the film to the solution. This is solid-state ionic conduction and depends on the presence of vacancies in a crystalline lattice. An amorphous lattice has no regular vacancies as does a solid crystal, and hence the rate of electrodiffusion of Fe through it is greatly diminished, as is the breakdown of the passive layer.
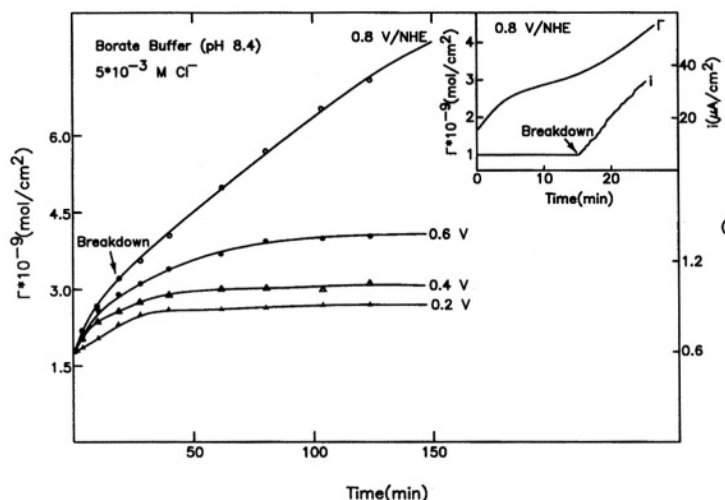
Fig. 12.67. Total "surface" concentration of chloride for 5×10^{-3} M Cl^- in the solution and different potentials as a function of time. Γ–t and i–t curves for $V = 0.80$ vs. NHE (insert). (Reprinted from V. Jovancicevic, J. O'M. Bockris, J. L. Canbajal, P. Zelenay, and T. Mizuno, *J. Electrochem. Soc.* **133**: 2223, 1986, Fig. 8. Reproduced by permission of The Electrochemical Society.)

concentrations of Cl^- in solution up to around 0.05 V, adsorption of Cl^- on the passive layer is related to time in the normal Langmuirian way (Fig. 12.67). However, as the potential is made more anodic, or the concentration of Cl^- is increased, there comes a point at which the surface concentration of Cl^- as a function of time changes its character. It goes above a monolayer in concentration and does not flatten out (i.e., approach a limit, one monolayer) with time, but continues to aggregate adsorbed Cl^- to amounts well above a monolayer. Thus, Cl^- is being *absorbed* into the passive layer, very direct evidence (Mizuno, 1986) for the basis of the depassivation model presented here.

12.4.6. Effects of Marine Organisms on Passive Layers

Some of the work on the breakdown of passive layers has been carried out in a marine environment, which is particularly important for the protection of ships. Although the main actor in passivity breakdown is still Cl^-, the situation is made more complex by the buildup of biofilms (films of dead marine bacteria) of great complexity on the metals concerned.

The potential of a free-floating stainless steel specimen over time (100 days) in a marine environment is driven in a positive direction ("enoblement"). To explain an
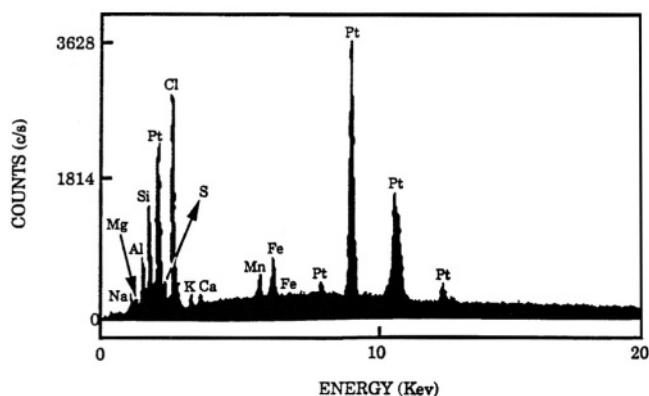
Fig. 12.68. EDAX data for a biofilmed platinum electrode exposed in sea water for 5 months. (Reprinted from P. Chandrasekaran and S. C. Dexter, NACE Conference paper 276, presented at Corrosion/94 with permission from NACE International.)

enhanced corrosion under these circumstances, it is necessary to account for an increased counter-cathodic current. Lewandorski and, independently Linhardt, have shown (1998) that bacteria produce MnO_2 in the films and it is the reduction of this MnO_2 that provides the enhanced cathodic current and speedier corrosion. Figure 12.68 shows the complex composition of the films after 5 months.

12.5. LOCALIZED CORROSION

12.5.1. Introduction

In 1939 Wagner and Traud used a flat and featureless plate to portray the basic theory of corrosion for a reason; such extreme simplification of the modeled system, made it possible to present a statement of the basic idea of why corrosion occurs.[17] However, real, actual corrosion, corrosion in technology, in moist environments everywhere, does not occur uniformly over the entire area of flat plates. It is nearly

[17]Previous to the Wagner and Traud paper of 1939, corrosion scientists imagined that electrochemical corrosion must always involve some kind of inclusion, some impurity atoms. In the first half of the century, electrochemists were extremely keen on the Electrochemical Series (Sec. 6.3.13.3) and the "thermodynamic driving force," by which they meant to the ΔG that would arise between the body of the corroding metal (M, anodic) and the impurity (M^1, cathodic). If the difference in the standard potential of these materials was ΔE, ($\Delta G = -nF\Delta E$), the local currents that are corrosion could be explained. What was not explained in the older theory was the surprising fact that corrosion of pure metal occurred, also. However, as the older workers slyly pointed out, what metal is *completely* pure?

always local (i.e., confined to a small area), and readers of this chapter have had a chance to learn something about that from the practical examples given at diagram level in Figs. 12.25–12.32. Local corrosion is the bane of industry and because it is of such great practical importance everywhere—the determiner of the practical life and therefore the economics of so many things—it is important that we learn something more about it.

12.5.2. The Initiation Mechanisms

At the most basic thermodynamic level, the driving force in a corrosion reaction occurs whenever two points in a system (one of which must involve the underlying solid) differ in their electrochemical potentials. Such situations occur repeatedly everywhere throughout nature, and the symbol $^{\alpha}\Delta^{\beta}\bar{\mu}$ for the difference of the electrochemical potential between point β and point α is a very compact way of indicating the driving energy[18] of all material happenings. Thus, it is pointless to try to list all the opportunities for the initiation of localized corrosion, but it may be helpful to give three examples.

12.5.2.1. *Forming a Pit or Crevice.* One mechanism for the breakdown of a passive layer was described in some detail in a discussion of how it could be possible for only 5% of a transition metal in Al to cut the corrosion rate of Al alloyed with it 100 times (Sec. 12.3.2). Once an invading particle starts to diffuse into the protective oxide, there is often a driving force for localized corrosion in the difference in the concentration of O_2 at the region of Cl^- entry, where O_2 is freely available, and the bottom of the pit, where it tends to be sparse (Fig. 12.27). Now, when O_2 is very available at one part of the system but little available at another, there tends to occur at this region of sparse oxygen availability a substitutive anodic reaction because then there is a substantial $^{\alpha}\Delta^{\beta}\bar{\mu}$ between the two ends. Thus, deep in the pit is a $\bar{\mu}_{M^{2+}}$ and at the top a $\bar{\mu}_{O_2}$. Pit growth is then often a consequence of this difference of partial molar electrochemical potentials, the "driving force."

12.5.2.2. *A Clamp on a Plain Piece of Metal.* Before the clamp (or some equivalent cover, see Fig. 12.31) is placed on a metal sheet immersed, say, in saline water in contact with air, the $\bar{\mu}_{O_2}$ is the same everywhere along the surface of the sheet of metal and $\Delta\bar{\mu}$ is zero, so no electrochemical reaction—no corrosion—occurs. However, the clamp prevents access of O_2 to the metal under it and therefore $\Delta\bar{\mu}$ becomes finite. Around the clamp there is usually a large area that is excellently aerated and able to supply electrons for the reduction of O_2. Of course, the electrons to do this have to come from somewhere in the underlying metal. In such a system, there is no

---

[18]The electrochemical potential, $\bar{\mu}_i^{\alpha}$ of a system is a partial molar free energy of the species, i, in the phase α together with an electrical energy term, $nF\phi^{\alpha}$ [λ is the inner potential of the phase, α (Chapter 6)]. $^{\alpha}\Delta^{\beta}\bar{\mu}_i$ is, then, the difference of two *electrochemical* potentials, or because of the meaning of the term "chemical potential," the difference of two electrochemical partial molar free energies. However, in terms of historical usage, $^{\alpha}\Delta^{\beta}\bar{\mu}$ is called (wrongly) "the driving force" of any reaction concerned.

external electrical power source, and a galvanic couple forms between the metal under the clamp, which kicks back electrons into the metal from the formation of ions in the moisture, and the $\bar{\mu}_{O_2}$ outside the metal surrounding the clamp. But when the O_2 from the surrounding air takes up electrons from the metal and gets reduced in water, the metal atoms underneath the clamp (no O_2) are embraced by the substantial $\Delta\bar{\mu}$ between the $\bar{\mu}_{O_2}$ at the bare metal and that of themselves and get persuaded to give up electrons and become ions, i.e., to dissolve, corrode. This example illustrates well the limitation on the life of practical objects set by local corrosion.[19]

Thus, local corrosion (and the term "local" may imply a size of a few atoms up to that of a millimeter) occurs whenever a region of a material surface, α, is connected electrically (through a flow of electrons in the underlying metal region, β, at which there are interfacial reactions exhibiting an electrochemical potential different from that at α. The different constituents of an alloy would tend to provoke such a situation; or, for example, S^- inclusions in steel.

12.5.2.3. *Pits in Stainless Steel.* Much work has been carried out on pits in stainless steel because of their particular technological importance. As with the Al case described in some detail in Section 12.3, the work on stainless steel in contact with Cl^- containing media points to an entry point for Cl^- diffusion through the passive film. These entry points (the beginning of pits) may arise from inclusions in the steel and for 304 stainless steel in particular, it is MnS inclusions that are critical (Szklarzcka-Schmialowska, 1986).

The use of photoresists to cover most of the metal surface and isolate individual pits has made possible detailed analysis of these entities, varying in radius from around 0.1 to 5 µm. The methods of examination include Auger spectroscopy, scanning electron microscopy, X-ray dispersive analysis, and atomic force microscopy (Ke and Alkyre, 1995).

Among the interesting results of such endeavors for 304 stainless steel is a stress upon inclusions involving manganese sulfide and manganese oxide. Surprisingly, traces of Cu in the steel are detected and form flowerlike deposits prior to the initiation of a pit. Metallic inclusions tend to dissolve anodically in 0.1 M NaCl at 400 mV SCE. There is some evidence for their readsorption. However, to be the cause of pit initiation, the inclusions have to be at least 0.7 μm in size. It may be that the local current density for dissolution at smaller inclusions is too high to sustain Cl^- adsorption there.

---

[19]The lethal character of much local corrosion in moist air depends on what kind of surrounding atmospheric conditions are present. Thus, air over the desert and air over the sea varies greatly in moisture content. Further, the conductivity of the moisture (solution) under the clamp varies as a function of the acidity of the rain, i.e., the SO_2 content of the air, which will depend on automobile and power plant emissions. Correspondingly, metal exposed to salt spray within a few miles of the shore will corrode more quickly than metal further away. The CO_2 content of the air provides the moisture, with H^+ and HCO_3^- ions for conductance, and increases with the growing CO_2 content caused by the use of fossil fuels.

12.5.3. Events in Pits

The initiation of pitting corrosion starts, then, with some kind of local irregularity containing metallic inclusions and continues with the penetration of Cl^- (or other aggressive ion) into the protecting layer at this point. Until the work of Brown (1970), the idea of finding out what happened inside a pit seemed impractical. However, Brown's finding[20] that the pH inside pits is independent of the pH in the external solution encouraged investigations in that direction which have been furthered by the development of ultramicroelectrodes (Section 7.5.4).

Pits tend to dissolve anodically at the bottom, and the resulting high concentration of cations near the crack tip may lead to an exceeding of the solubility product, e.g., of $Fe(OH)_2$ and the formation of blocking deposits which thus increase the ionic resistance in the pit. The effect of the ohmic drop inside the pit seems to be more important than earlier analysts (Cherapinov, 1986) had thought. It increases the potential difference between top and bottom of a pit. Since the electrochemical situation at the top is dominated by the cathodic reduction of oxygen (occurring at a local potential of about –0.1 V on the normal hydrogen scale), the IR drop tends to increase the potential difference between top and bottom and help create a condition for metal (e.g., Fe) dissolution to occur at a local potential of about –0.2 on the same scale. The condition at the bottom of crack spots more will be discussed in more detail in Section 12.6.4 on stress corrosion cracking.

Momarcky interference microscopy has been used to look at pits revealed in various depths by sputtering (Proost, 1998). Its images show fractal patterns on the sides of the pits.

12.5.4. Modeling

Although there has been an increase in the ability to look inside pits with the aid of ultramicroelectrodes and find out directly what happens there, there is considerable counter-attraction for a mathematical simulation of these events. The objective of such modeling (White, 1990) is not primarily to obtain a correct theory at a molecular level of what is happening. It is, rather, to discover equations (algorithms) that work, i.e., that allow the effect of certain variables to be predicted. Alkire (1990), has used mathematical simulations to investigate a complication neglected in molecular-level theories. It is the effect of the external flow of a solution. Corrosion may generate dissolved products that may either diffuse away from the surface or be readsorbed; in the latter case, the products may lead to sites on the surface that may initiate pits. A fast-flowing external solution would tend to remove the dissolved products and prevent readsorption.

---

[20]Brown and his colleagues constructed an apparatus in which a solution traveled the length of an artificially contrived crack. The solution was made to wet a filter paper and its pH was measured by means of a glass electrode with a long, thin, tip.

Correspondingly (Pickering, 1995) the IR drop mentioned above can be modeled, and in crevices and pits, the model can lead to a calculation of the depth below the orifice (where the cathodic reduction of oxygen is dominant) at which the potential has been sufficiently changed so that active dissolution begins. Lillard and Scully (1994) modeled not only the IR drop but also the change of composition in crevice corrosion; the local corrosion currents were modeled with the help of an equivalent circuit approach (Section). Remarkably explicit and detailed (though experimentally unconfirmed) results from such analyses are obtained at far less a cost and time than that of an experimental approach that is sometimes not feasible. An example from Scully's work is shown in Fig. 12.69. It shows the complex variation of the rate and place of anodic dissolution in a crevice, together with the variation of the magnitude and position of the cathodic partner current. Now, indeed, one is very far from Wagner and Traud's featureless corroding plane!
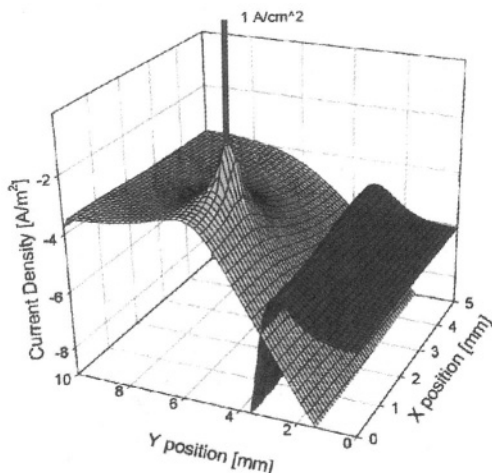
Fig. 12.69. Current distribution resulting from polarization of mouth of a crevice to +0.3 V (SCE). Dark represents anodic current, light represents cathodic current. (Reprinted from R. G. Kelly, "Measurement and Modeling of Localized Corrosion Sites," *Interface* **6**(2): 21, 1997, Fig. 5c. Reproduced by permission of The Electrochemical Society.)

Further Reading

Seminal

1. M. Faraday, in *Experimental Researches in Electricity,* Vol. 2, London, 1844; Reprinted by Dover, New York (1965). First suggestion of passivity as due to thin film.
2. N. Cabrera and N. F. Mott, *Rep. Prog. Phys.* **12**: 163 (1943). Theory of the state of growth of oxide films.
3. U. R. Evans, *Trans. Electrochem. Soc.* **91**: 547 (1947). Isolation of a film from a passive metal by dissolving away the metal.
4. C. Wagner, *J. Electrochem. Soc.* **99**: 369 (1952). Alloying with noble metals to protect corroding metals.
5. T. P. Hoar and J. G. Hynes, *J. Iron Steel Inst.* **182**: 124 (1956). Time to failure in alloys.
6. M. J. Pryor, *J. Electrochem. Soc.* **106**: 557 (1959). First statement of Cl^- penetration theory of depassivation.
7. C. Edeleanu, *Chem. Ind.* **301**: 50 (1961). Why passive films remain constant in thickness during variation in potential.
8. A. K. Reddy and J. O'M. Bockris, *J. Bur. Standards,* p.229 (1964). First ellipsometric observation of passive films on electrodes in solution under potential control.
9. H. Pickering and C. Wagner, *J. Electrochem. Soc.* **114**: 698 (1967). Paired vacancy diffusion in alloy dissolution.
10. H. H. Uhlig, *Corros. Sci.* **7**: 235 (1967). First statement of idea that passivity is due to a *monolayer* of oxide.
11. B. F. Brown, *J. Electrochem. Soc.* **116**:218 (1969). First measurement of pH inside pits.
12. J. O'M. Bockris, M. Genshaw and V. Brusic, *Symp. Faraday Soc.* **6**: 177 (1970). Comprehensive application of ellipsometry to Fe passivation.
13. W. E. O'Grady and J. O'M. Bockris, *Chem. Phys. Lett.* **5***:* 116 (1970); *Surf. Sci.* **66**: 581 (1977). First application of Mössbauer spectroscopy in electrochemistry; the properties of passive films are due to their amorphous character.
14. J. O'M. Bockris, B. T. Rubin, A. Despic, and B. Lovrecech, *Electrochim. Acta* **17**: 97 (1972). Cu-Ni alloy dissolution; the dissolution rate of each alloying component is independent of its composition in the alloys.
15. J. Gniewich, J. Pezy, B. G. Baker, and J. O'M. Bockris, *J. Electrochem. Soc.* **125**: 17 (1978). First Auger study of the surface of a dissolving alloy (attempt to show that small concentrations of gold would protect less noble metals).

Reviews

1. J. W. Schultze and S. Kudeka, "Investigation of Passivity," *Interface* **6**: 28 (1997).
2. P. Schmuki and S. Virtanen, "Modeling of Passivity," *Interface* **6**: 38 (1997).
3. R. G. Kelly, "Small-Scale Corrosion," *Interface* **6**: 18 (1997).

Modern

1. R. W. Revie, B. G. Baker, and J. O'M. Bockris, *Surf. Sci.* **52**: 664 (1975). First published paper on uhv in electrochemistry; the degraded passive film by Auger spectroscopy.

2. O. J. Murphy, T. E. Pou, J. O'M. Bockris, L. L. Tongsen, and M. D. Monkowski, *J. Electrochem. Soc.* **130**: 1792 (1983). Water in the passive layer; SIMS and ISS evidence.

3. P. M. Natashan, E. McCafferty, and G. K. Hubler, *J. Electrochem. Soc.* **133**: 1061 (1986). pH, pzc, and the corrosion of Al alloys.

4. R. Alkire and K. P. Wong, *Corr. Sci.* **28**:411 (1988). Microelectrodes in pitting corrosion.

5. B. F. Shew, G. D. Davis, T. L. Fritz, B. J. Rees, and W. C. Moshier, *J. Electrochem. Soc.* **138:** 3288 (1991). Enrichment in the surface of Al alloys.

6. Z. Szarkloska-Schmialowska, *Corros. Sci.* **33**: 1193 (1992). A solution in pits has a pH that allows dissolution of metal oxides.

7. J. O'M. Bockris and L. Minevski, *J. Electroanal. Chem.* **349**: 375 (1993). Protection of aluminum by means of transition metal alloys.

8. N. Casillas, S. J. Charlebois, W. H. Smyrl, and H. S. White, *J. Electrochem. Soc.* **140**: L142 (1993). Confocal laser scanning microscopy used on electrodes.

9. R. Raiceff, I. Betova, M. Bohnov, and E. Lazarova, in *Modeling Corrosion,* K. Trethoway and P. Roberge, eds., Kluwer Academic, Dordrecht (1994). Modeling of corrosion reactions.

10. G. Salamat, G. Juhl, and R. G. Kelly, *Corrosion* **51**: 826 (1995). Local concentrations are significantly different from bulk ones.

11. H. S. Isaacs, S. M. Huang, and V. Jovancicevic, *J. Electrochem. Soc.* **143**: 1178 (1996). Impedance measurements in pits.

12. A. Michaelis and J. W. Schultze, *Thin Solid Films* **274**: 82 (1996). Photoelectrochemical examination of passive layers.

13. J. O. Park, C. H. Park, and R. C. Alkire, *J. Electrochem. Soc.* **143**: L174 (1996). Microelectrodes in corrosion research.

14. J. O'M. Bockris and Y. Kang, *J. Solid State Electrochem.* **1**: 17 (1997). Potential of zero charge and the protection of aluminum from Cl$^-$ attack by transition metal additives.

15. F. Mansfeld, G. Zhong, and C. Chen, *Plating Surf. Finishing* (Dec.) 72 (1997). Impedance measurements on aluminum.

16. J. Proost, J. Baklanov, M. Verbeeck, and K. Mrex, *J. Solid State Electrochem.* **2**: 150 (1998). Looking inside pits.

12.6. ELECTROCHEMICAL ASPECTS OF THE EFFECT OF HYDROGEN ON METAL

12.6.1. Hydrogen Diffusion into a Metal

It has been sufficiently emphasized that the *instability of metal surfaces arises from an electrodic mechanism*; an electronation reaction teams up with the metal-dissolution reaction to keep numerous micro *corrosion cells* running.

What has all this to do with the inside of the metal? One would think that the inside is sufficiently isolated from the surface to remain safe and stable. It will be shown, however, that events at the borders of a metal may have internal repercussions

and eventually cause even the inside of the metal to decay, i.e., to lose its mechanical properties. Thus, electrodic charge-transfer reactions at the surface have far-reaching implications for the strength of bulk metals. It is intended here to indicate briefly how the surface instability can be propagated to the inside of the metal.

Consider a corroding metal, and let hydrogen evolution be the electronation reaction. The formation of hydrogen atoms adsorbed on the metal surface is an essential intermediate step in the electrodic evolution of hydrogen. What happens to these adsorbed hydrogen atoms? They can get desorbed in either a chemical or electrodic reaction as hydrogen molecules that diffuse out into the solution or collect in bubbles of hydrogen gas. This is the visible *way out* from the metal surface.

But there is also a *way in* from the surface; the *adsorbed* hydrogen atoms can dissolve into the metal to form *adsorbed* hydrogen. Since one has started off with zero concentration of adsorbed hydrogen inside the bulk of the metal, a gradient of hydrogen concentration develops between the surface where hydrogen enters and the interior of the metal. This concentration gradient makes the adsorbed hydrogen diffuse into the metal. The extent of this phenomenon is exemplified in Table 12.4 for different types of iron. The diffusion coefficients are seen to be of the same order of magnitude as those for diffusion of ions in aqueous solution. Thus, diffusion of hydrogen in the bulk of the metal can be considered a fairly fast process.

The permeation of hydrogen into the interior of a metal can be shown in a simple way (Fig. 12.70). A thin, metal membrane separates two vessels containing an electrolyte. Hydrogen is electrodically evolved on one side. The potential difference across the other membrane-electrolyte interface is adjusted for the deelectronation (or ionization) of any adsorbed hydrogen $H_{ads} \rightarrow H^+ + e$, coming through the metal from the entry side of the membrane. Thus, the ionization current is the manifestation of the hydrogen permeating through the metal membrane from the surface on which hydrogen is evolved. Quantitative correlations can be made, as shown below, between the permeation current and the diffusion coefficient and the flux of hydrogen through the metal.

The steady-state deelectronation current density i_d is related by the condition of flux equality (Section 4.2.7) to the flux J_p of hydrogen permeating through the metal (Fig. 12.70):

TABLE 12.4
Diffusion Coefficients and Critical Concentrations of Hydrogen for Diffusion in Iron at 25 °C

Material	D_K (10^{-5} cm^2 s^{-1})	c_k (10^{-8} mol cm^{-3} Fe)
Polycrystalline Armco iron	6.3	14.3
Single-crystal iron	8.3	12.0
Zone-refined iron	6.1	13.9
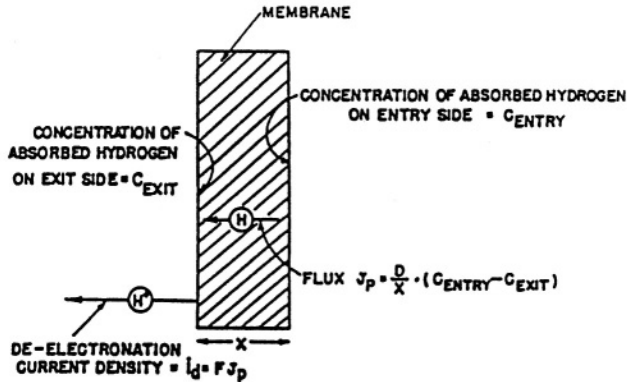
Fig. 12.70. The penetration of hydrogen into the bulk of a metal can be detected if the metal is sliced into a membrane and anodic dissolution of hydrogen is observed on the side opposite to that at which the evolution takes place.

$$\frac{i_d}{F} = J_p \tag{12.56}$$

and, using Fick's law for the steady-state permeation flus, one has

$$\frac{i_d}{F} = J_p = \frac{D(c_{entry} - c_{exit})}{l} \tag{12.57}$$

where c_{entry} and c_{exit} are the concentrations of absorbed hydrogen on the entry and exit sides of the membrane of thickness l. The potential difference across the membrane-electrolyte on the exit side can be adjusted so that $c_{exit} = 0$, i.e., all the hydrogen coming through is immediately ionized. Thus,

$$\frac{i_d}{F} = \frac{Dc_{entry}}{l} \tag{12.58}$$

Hence, by measuring the permeation current, it is possible to study the diffusion coefficient of the hydrogen inside the metal.

12.6.2. The Preferential Diffusion of Absorbed Hydrogen to Regions of Stress in a Metal

In a polycrystalline material, one might have considered that the grain boundaries are *irrigation channels* for the diffusing hydrogen. Yet, it is found that the diffusion coefficient for hydrogen is the same for polycrystalline and single-crystal iron. Thus, the hydrogen must be diffusing through the lattice—interstitial diffusion (Fig. 12.71).
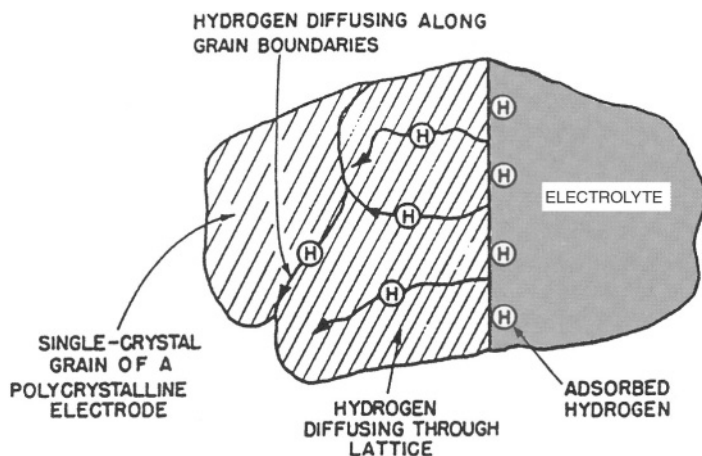
Fig. 12.71. Hydrogen atoms adsorbed at the surface may diffuse into the bulk of the metal by intercrystalline, or interstitial, diffusion (diffusion *through* the crystal lattice).

When a hydrogen atom occupies an interstitial site, there is a certain displacement of the atoms around the interstitial atom; it is as though some atoms moved apart a little to accommodate the interstitial hydrogen atom (Fig. 12.72). Hence, one may consider that there is a certain change in volume due to the entry of a hydrogen atom into the lattice. The net change in volume resulting from 1 g-atom of hydrogen is the partial molar volume $\bar{V}_H$, of hydrogen in the metal.

It is possible to obtain the partial molar volume of hydrogen in a metal provided one knows the solubility of hydrogen in it, corresponding to a constant pressure or overpotential, *as a function of applied stress.* For the applied stress to be thermodynamically significant, it should be within the Hooke's-law region for the metal (Fig. 12.73). Proceeding from the thermodynamic relations $(\partial\mu/\partial P)_T = \bar{V}$ and $\mu = \mu^0 + RT \ln c$ (when c is small), one has:

$$\left[\frac{\partial \ln(c/c_0)}{\partial P}\right]_T = \frac{\bar{V}_H}{RT} \tag{12.59}$$

Here c is the solubility of hydrogen (g-atom H/cm^3 metal) when the applied uniaxial stress (equivalent to pressure, see below) is σ and c_0 is the solubility when the applied stress is zero. The general relation between the applied stress and pressure or the equivalent hydrostatic stress σ_h, can be written as

$$P = \tfrac{1}{3}(\sigma_x + \sigma_y + \sigma_z) = \sigma_h \tag{12.60}$$
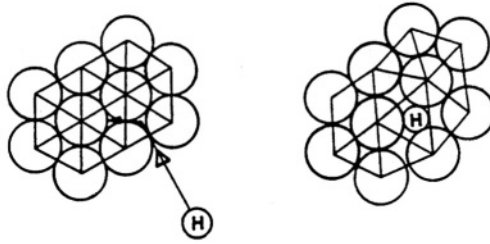
Fig. 12.72. As a hydrogen atom penetrates into a lattice, a slight distortion takes place to accommodate it.

where σ_x, σ_y, and σ_z, are the components of the applied stress. For uniaxial stress condition, $\sigma_x = \sigma_z = 0$ and hence, $P = \frac{1}{3}\sigma_y = \frac{1}{3}\sigma = \sigma_h$. Substituting for P in terms of σ_h (negative for tensile stress) in Eq. (12.59) and integrating, one obtains

$$c = c_0 e^{\sigma_h \overline{V}_H / RT} \tag{12.61}$$

Measurements of c and c_0 (the hydrogen solubility in the presence and absence of stress) can be made by determining the permeation of H at a series of stresses. Equation 12.61 was used by Beck, Nanis, and McBreen to determine $\overline{V}_H$ for H in pure
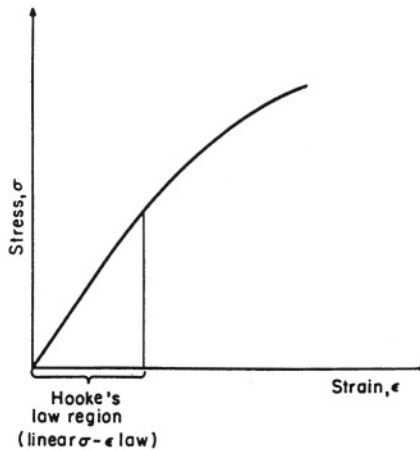
Fig. 12.73. Hooke's law relates the strain produced in a body under stress (the linear stress-strain laws). Past a certain stress, the law is no longer obeyed and a plastic deformation of the metal takes place.

iron. Their value (numerically corrected for the absence of a $(2.303)^2$ factor) is 2.6 cm^2/mol.

Thus, from Eq. (12.61) and taking σ_h as positive when the stress is tensile and negative when it is compressive, one sees that hydrogen accumulates in regions of compressive stress. The stress need not be externally applied.

Equation (12.61) is also true for residual stresses in metal. The latter kind of stress usually will have tensile *and* compressive stress fields associated with it. As far as the solubility of hydrogen is concerned, the effect of the tensile stress field (which *increases* solubility) overwhelms the counter-effect due to the compressive stress field (which tends only to decrease the already small solubility). Therefore, the larger the lattice strain or distortion, the larger the concentration of hydrogen (Fig. 12.74). All imperfections in crystals are regions of distortion or strain. Hence, absorbed hydrogen finds its way to, and concentrates at, such imperfections.

There is a simple way of experimentally demonstrating that stretching a metal leads to increased hydrogen absorption (McBreen, 1976). The permeation currents measured by the membrane technique should show an increase when the membrane is stretched. This is, in fact, what happens (Fig. 12.75).

12.6.3. Hydrogen Can Crack Open a Metal Surface

It is known that the imperfections in a metal include voids that are larger than atomic dimensions, say about 100 Å across. On reaching these regions, the *ab*sorbed hydrogen atoms feel they have reached an exposed surface. They become *ad*sorbed hydrogen atoms and combine to form hydrogen molecules: a chemical desorption
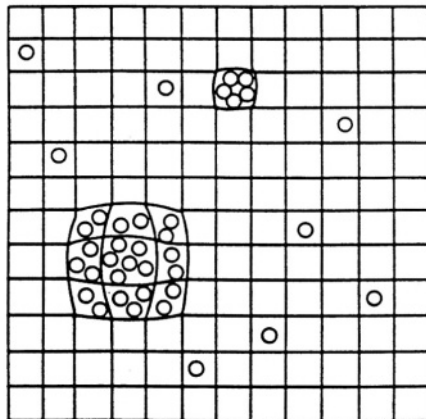
Fig. 12.74. Hydrogen atoms accumulate at those points in the lattice at which there are larger interstices produced by stretching the specimen.
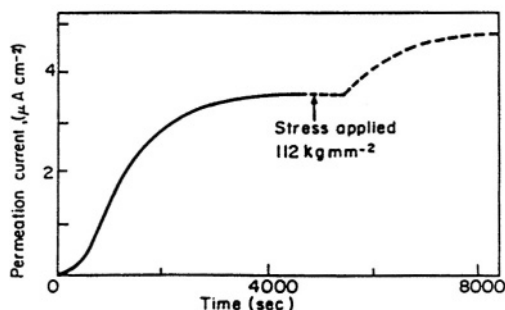
Fig. 12.75. Hydrogen atoms absorbed in the metal recombine when they reach the voids and form molecular hydrogen inside the metal.

reaction takes place, i.e., $2H_{ads} \rightarrow H_2$ (Fig. 12.76). Thus, a pressure of hydrogen gas builds up inside the void. Calculations show that the pressures can amount to thousands of atmospheres, indeed a pressure so high that the surrounding metal is stressed beyond its elastic limit. The metal yields and the void becomes a cavity (supervoid) as discussed in Section 12.6.6.

If all this happens near the surface of the metal from which hydrogen is entering the metal, a blister may be formed near the surface. Eventually, the walls of the blister collapse and this rupture allows the gas to escape. In the process, however, a crack has been initiated at the metal surface (Fig. 12.77). The whole process of crack initiation is of course facilitated if an outside stress is applied to the metal. Then, if the metal structure is such that there are specially high stresses at some points, it is precisely at these points that there is the greatest likelihood of crack initiation because hydrogen permeates preferentially into the stressed region and enters the voids nearest the stressed region. The cracks thus *initiate* there.
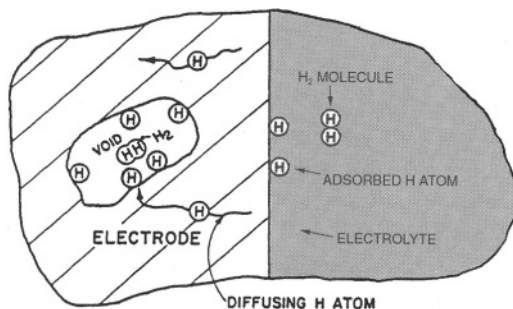
Fig. 12.76. Hydrogen atoms absorbed in the metal recombine when they reach the voids and form molecular hydrogen inside the metal.
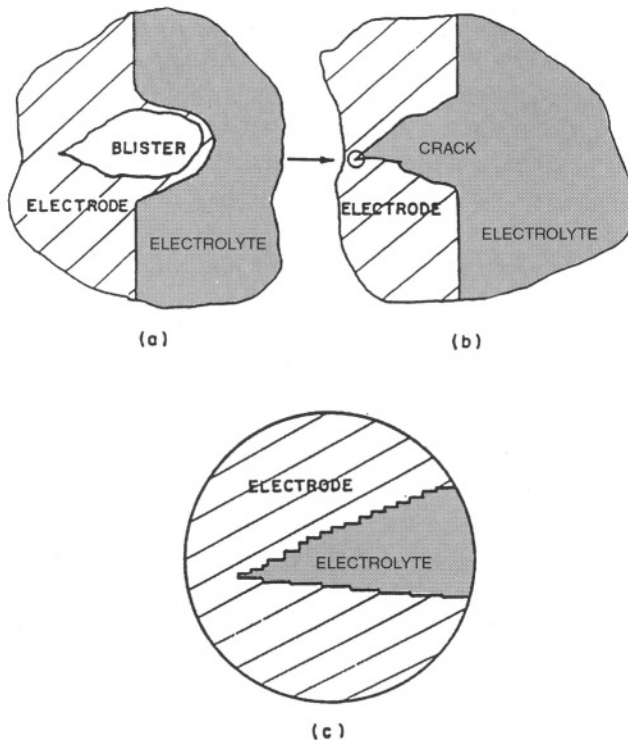
Fig. 12.77. The metal near the surface yields under the pressure of the accumulated molecular hydrogen, forming a blister (a), which breaks up and initiates a crack in the metal (b). At the bottom of the crack (c), the stress is very great and may produce further yielding.

Another mechanism for the initiation of cracks in a surface innocent of adsorbed aggressive anions but containing adsorbed H depends on the surface activity of H. Thus, at points of defect on the surface (e.g., near emergent screw dislocations), the surface stress is abnormally high (Subramanyam, 1986) and for this reason, and following Eq. (12.61), the local solubility will be abnormally high. At sufficiently high local concentrations of H, there will be a weakening of M–M bonds, and the eventual local breakdown of the strength of the surface will give rise to opportunities for rapid ingress of H at high fugacities.[21]

---

[21]Hydrogen initiation of stress-corrosion cracking is indeed the probable mechanism. However, what has been given here is rather overgeneral. For example, the stress corrosion of alloys shows specificities that hint at unexplained factors. Passive films form at the bottom of pits and it is the breaking of these upon stress that sometimes causes cracks to spread.

12.6.4. Surface Instability and the Internal Decay of Metals: Stress-Corrosion Cracking

So far, corrosion has not come into the picture except as it stimulates the $H^+ + e \rightarrow H_{ads}$ electronation reaction and is therefore responsible for the hydrogen accumulation inside the metal. Now consider a metal that is simultaneously being corroded and some parts of which are subjected to a tensile stress. The permeating hydrogen will tend to initiate a crack in the region where the stress is great by the mechanism described in the previous section, and the electrolytic solution (the corrosive environment) comes into contact with the inside of the crack (Fig. 12.77).

Once the crack is initiated, the metal surface inside the crack may be quite different from the normal surface of the metal. Thus, in the course of plastic deformation, the metal could have developed slip steps [see Fig. 12.77(c)] which contain crystallographic planes of high Miller index at which the specific dissolution rate (or exchange current density) may be larger than that at the normal metal surface. Anodic current densities of some 10^4 times those at a passive surface have been shown to appear at a metal surface that is yielding under stress (Despic and Raicheff, 1978).

One conclusion is clear. The instability of a metal with surface cracks will tend to be greater than that of a surface without such cracks. The metal-dissolution and hydrogen-evolution reactions tend to occur indiscriminately on the normal surface of a homogeneous single crystal. When, however, there is a crack, the metal dissolution will occur preferentially inside the crack and the hydrogen evolution on the surface outside the crack (Fig. 12.78). But this implies that the electron-source area A_H is very large compared with the area A_M inside the crack, i.e., compared with the area over which there is metal dissolution. It is essential, however, that the corrosion current (not the current *density*) be equal to the electronation current:

$$I_{corr} = I_M = |I_H| \tag{12.62}$$

$$= A_M i_M = A_H |i_H| \tag{12.63}$$

Hence,

$$i_M = \frac{A_H}{A_M} |i_H| \tag{12.64}$$

and, since $A_H \gg A_M$, it follows that

$$i_M \gg |i_H| \tag{12.65}$$

i.e., even though the hydrogen-evolution current density is small (say, 10^{-4} A cm^{-2}, the metal-dissolution current density may be greater by a factor of A_H/A_M, which can be on the order of 10^3. Such high current densities can be sustained by the
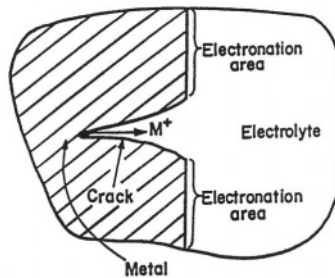
Fig. 12.78. When a stress crack appears at the surface, this becomes a *locus of an increased* dissolution activity, the electrons being drawn away to the rest of the surface as the electronation area.

metal-dissolution reaction inside the crack because of the abnormally high exchange current densities possible there.

A model has been given for high dissolution rates inside cracks in terms of the oxide-free and highly kinked state of the surface there. What happens when the metastable kinks are dissolved off or when a passive film is formed to cover the walls of the crack? The surface inside the crack becomes more normal, and so does the current density in that region. Thus, the dissolution rate should become normal inside the crack.

Now consider that the whole corroding metal is being stretched by a tensile stress. The gross average applied stress is not sufficient to make the metal yield; the stress is within the elastic limit. Release it, and the metal will spring back to its original dimensions. This does not mean, however, that the local stresses are equal to the average stress. An abnormally high stress concentration at the apex of the crack need not arise from externally applied stresses; the high stress may be due to residual stresses left behind in the metal at the time of its incorporation into a fabricated structure (e.g., the region around the rivets or welds in a steel boiler).

Irrespective of how the abnormal stress concentration arises, it is possible that the material at the crack apex is locally stressed into the plastic deformation region of the stress-strain curve. What is the result of the yielding of the metal near the crack apex? The result is that as anodic dissolution dissolves away the kinky surface, further plastic yielding creates a fresh kinky surface inside the crack, and thus the yielding helps the metal dissolution along at a rate (e.g., of millimeters per hour) that turns out to be far greater than what would be expected at the overpotential concerned from measurements on the normal surface (Hoar and West, 1976).

This is not all. What is happening is that the crack is propagating into the interior of the metal, with the advancing edge of the area serving as the electron-sink area for the metal-dissolution reaction. Superficially, everything is normal; if one measures the potential difference between the solution and the apparent surface of the metal, one gets almost the usual corrosion potential. Then microcracks begin to join up with other microcracks, macrocracks are produced, and the piece of metal ceases to be a stable structural material; an axle cracks, or a part of an aircraft disintegrates when it is only normally stressed.

What has been described is what is called *stress-corrosion cracking*. Some common examples of systems that tend to undergo this type of corrosion are given in Table 12.5. But perhaps one should call it *yield-assisted corrosion* (an electrochemical-plus-mechanical phenomenon) in contrast to normal *field-assisted dissolution* (an electrochemical phenomenon).

At this point, one may feel a lurking doubt. Maybe the crack-propagation process has nothing to do with electrodic dissolution, and the stress by itself does the damage. This view can be tested simply by superimposing a double-layer field and adjusting the metal-solution potential difference so that metal dissolution stops. Despite the continued presence of the stress, the crack propagation stops—clear evidence that for stress-corrosion cracking to occur, both the stress and the abnormally high corrosion rate inside the crack are essential. Without the stress, the dissolution rate inside the crack becomes normal and the crack ceases to advance into the metal; without the

TABLE 12.5
Some Common Examples of Stress-Corrosion Systems

Alloy	Medium	Type of Cracking
18 Cr–8 Ni	Cl^-	Transgranular
Steels	OH^-	Transgranular
70 Cu–30 Zn	NH_4^+ and some amines	Transgranular at low pH, intergranular in neutral solutions
Al–4% Cu alloys	Cl^-	In regions adjacent to grain boundaries
Al–7% Mg alloys	Cl^-	Intergranular
Magnesium alloys	Cl^-	Transgranular or intergranular
Titanium alloys	Cl^- Transgranular and intergranular	
Mild steel	NO_3^- and OH^-	Intergranular
Cu–P and Cu–Al	NH_4^+ and some amines	Intergranular
β brasses	Cl^-	Transgranular
	NH_4^+	Intergranular
Cu_3Au	$FeCl_3$	Intergranular

Source: Reprinted from J. C. Scully, *The Fundamentals of Corrosion*, Pergamon Press, 1966.

electrochemical dissolution at the crack apex, the stress cannot of itself make the crack advance into the body of the metal.[22]

This dual effect of the influence of stress upon the acceleration of the actual dissolution rate of a crack tip at a given overpotential (Despic and Raiceff, 1968) and the embrittling effects of H in the metal surrounding the tip are of interest in respect to the application of seminal equation (12.61). It is seen from this equation that if the local stress is pulling the metal apart and is negative, the local solubility of H will increase.[23] Hence, at the crack tip (a point of high tensile stress), the local concentration of H will be considerably higher than what it is when H_2 is in equilibrium with a metal at 1 atm. To know the degree by which this enhancement will take place, one will have to know the partial molar volume for H in the metal $(\overline{V}_H)$. This is $2.6 \text{ cm}^3 \text{ mol}^{-1}$ for H in pure Fe, but of course the value (a critical one for understanding damage) will be different for other metal alloys of Fe and for other metals. Then the local stress σ, at the crack tip must be known, also. The local H solubility at the crack tip in pure iron and that in its bulk of the metal have been calculated to be as much as 10^6 in favor of the tip. Now, the solubility of H in pure iron in terms of moles of H per mole of Fe is about 10^{-6} (at 1 atm of H_2) so that at the crack tip a local concentration of H equal to 1 atoms of H per 2 atom of Fe may occur. One can at once see the destructive consequences of this. M–H bonds will be formed and M–M ones will break, and the strength of the metal at the crack head will greatly weaken. Thus, the advance of a crack or pit in a metal will depend, not only on the enhanced $(i_0)_M$, an increased exchange current density for the dissolution of the metal there, but also on the damaged (hence weakened) nature of the metal at the crack tip.

All this is background for a description of a remarkable result found by microphotography of a spreading crack (Flitt, 1977). Metal bent in the form of a ring and surface stressed is placed in boiling $MgCl_2$ solution, and then observed under time-lapse photography and examined for cracks. After a short latency period a crack forms, quite reproducibly, and proceeds vertically down the metal constituting the ring (Fig. 12.79). Two important aspects of this crucial experiment may be recorded here.

1. Near the bottom of the spreading crack, a violent (cathodic) evolution of hydrogen is seen. This does not interfere with the notion that the actual crack tip is the site of an anodic metal dissolution. However, it is interesting that the hydrogen evolution occurs so near the crack tip, for in other surroundings (e.g., in moist air) one might have expected oxygen reduction to be the cathodic partner and occur over the ring's surface.

2. More interesting still are the characteristics of the motion of the crack, which can be seen if the time-lapse film is advanced at a speed to compress 1 hr of observations

---

[22]This is not to say of course that all cracking of materials has an electrochemical step. It is clear, e.g., that the cracking of glass (a nonconductor) is unconnected with anodic dissolution. If the energy associated with stressing a material is greater than the local surface free energy, spontaneous cracking will occur.

[23]Equation (12.61) is written to conform to current convention so that a negative σ (stress) *increases* solubility.
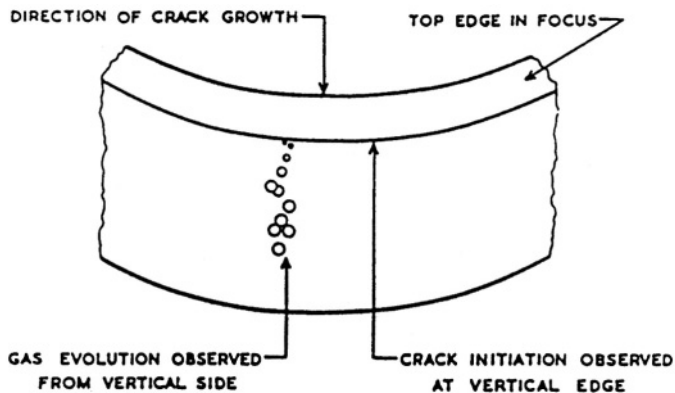
Fig. 12.79. Initiation site of a crack.

into 5 min. One sees the crack advancing at a slow rate (in the range of millimeters per hour) but then, every now and again the crack slips forward at a rapid rate. It seems likely that this latter movement is the result of embrittlement at the crack tip. The slow electrochemically controlled rate of spreading brings to the region around the crack tip an embrittled volume of metal because the high tensile stress at the tip of cracks causes a large increase in H solubility there and hence an area of weakened H embrittled metal [cf. Eq. (12.61)]. When the tip "feels" it has come upon this weakened section of metal (and since it is under local stress), it "tears" the metal apart and easily advances through the damaged region. After slipping thus through a length on the order of 0.1 μm, the tip is outside the embrittled region and resumes its stately movement through the metal under electrochemical corrosion control. However, it is now in new territory, undamaged metal, and for a little time continues by means of anodic dissolution with the cathodic hydrogen evolution partner. But of course a crack tip is under tensile stress and so the surrounding hydrogen in the metal bulk congregates there (increase of solubility with Fe inside stress) and makes a new damaged weak region, whereupon the tip of the crack advances rapidly once more through the weakened region, and so on (see Fig. 12.79). Hence, although it is electrochemical corrosion that causes some advance of the crack, the average velocity is dominated by *tearing* of the metal through its H-damaged region.

This combination of electrochemical and mechanical mechanisms, is the essential mechanism of stress-corrosion cracking. However, although the story has advanced, it is no use pretending that there are no more chapters to come. Indeed, Doig and Flewit (1995) have added further vital aspects in introducing the effect of slip planes into the consideration of the rate-determining step in the advance of a crack. The various crystal planes in a metal have different rates in their ability to slip and flow under stress. Not every plane will undergo the movements characteristic of cracking.

12.6.5. Practical Consequences of Stress-Corrosion Cracking

Quite dramatic things can and do happen as a result of cracks that develop as a result of corrosion. Bridges fall down. Several ships per year disappear, probably because they have suddenly split apart in a storm in which waves lifted the bow and stern of the vessel, leaving it momentarily suspended with too much tensile stress on the middle section, so that a crack appeared there and rapidly spread. Figure 12.80 shows typical examples of metal "ends" that have "broken" in stress corrosion cracking, an insidious danger for any structure under stress, particularly when, through corrosion reactions, surface-produced H can enter the metal and give rise to local embrittlement conditions.

12.6.6. Surface Instability and Internal Decay of Metals: Hydrogen Embrittlement

It has been known for quite some time that some very strong metals may suddenly lose their strength and become brittle even though there is no indication of an applied
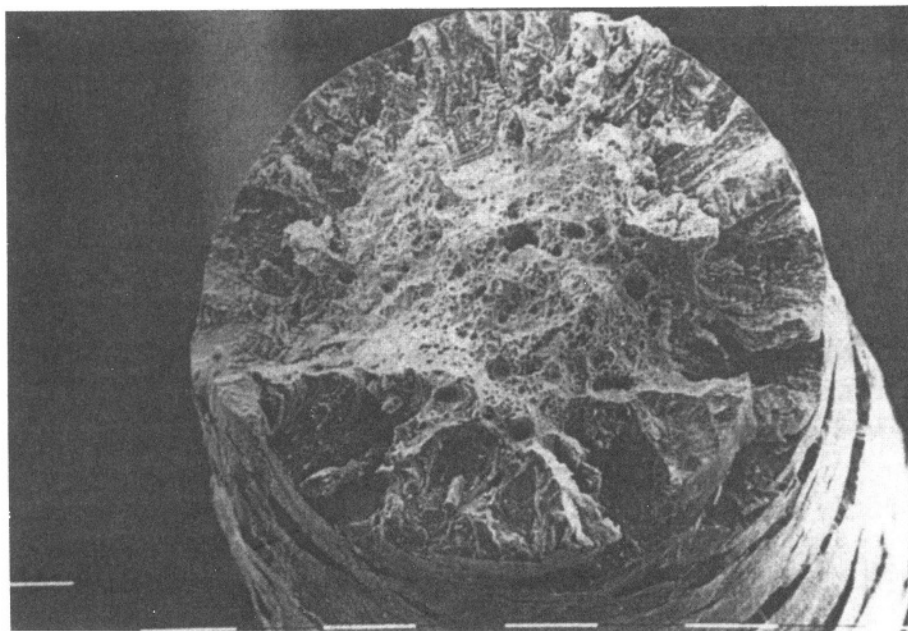
Fig. 12.80. Samples of a brass (64% Cu-26% Zn) alloy bar strained in a NaNO$_2$ solution at 0.2 V vs. NHE. Typical stress-corrosion cracking fractures (dashes = 100 μm). (Reprinted from J. R. Galvele, "Electrochemical Aspects of Stress Corrosion Cracking," in *Modern Aspects of Electrochemistry*, R. E. White, J. O'M. Bockris, and B. E. Conway, eds., No. 27, p. 234, Plenum, 1995.)

or initial stress. Thus, a copper wire exposed to H under certain circumstances can become so brittle that it can easily be broken apart in the hands like paper; the metal has become embrittled.

What is the mechanism of this phenomenon? Very early during investigations of this field, it was realized that metals become embrittled because at some stage of their career, their surface was the scene of a hydrogen-evolution reaction either because the metal was deliberately used as an electron-source electrode in a substance-producing cell or because parts of the metal became electron-source areas in a corrosion process. In fact, the phenomenon has come to be known as *hydrogen embrittlement.*

An approximate picture of what happens during hydrogen embrittlement can be sketched. The process commences with hydrogen's diffusing into the metal and accumulating in distorted regions of the lattice. Any voids (tiny cavities) in the lattice permit the accumulation of hydrogen gas by the chemical desorption of hydrogen atoms (supplied by diffusion from the surface). If the pressure of the gas become sufficiently high, cracks or large cavities are initiated. The result of all these events is that the metal ends up with plenty of cracks inside it (Fig. 12.81). When stretched, it
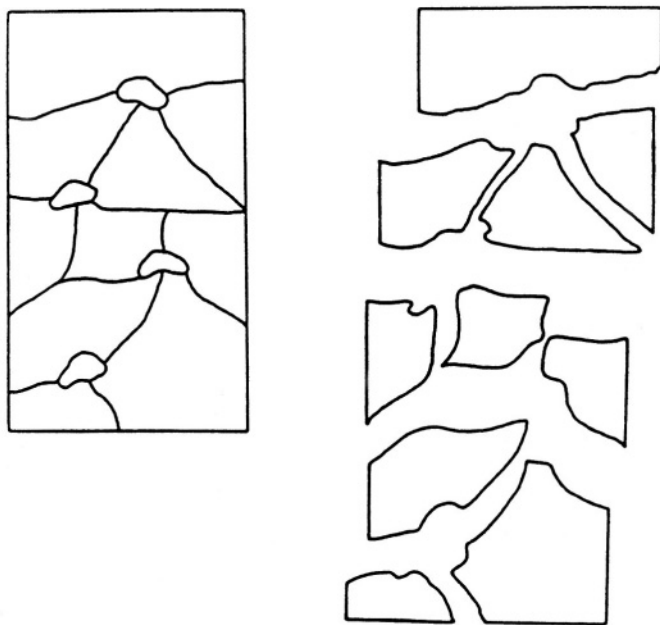
Fig. 12.81. Accumulated hydrogen (originating in electrochemical formation at the surface) produces cracks inside a metal. When it is strained, the metal fractures along the cracks rather than yielding. It has become brittle.

does not yield like a ductile material; it fractures along the cracks. The hydrogen has embrittled the metal.

It is possible to write down approximate conditions for crack propagation in a very simple way. Consider the *Griffith crack*, as it is called, a disk-shaped crack with a length l (Fig. 12.82).

The strain energy U_{strain} is the work done to increase the strain of the crack from $\varepsilon = 0$ to $\varepsilon = \varepsilon$ under the action of the stress σ that arises from the pressure of the hydrogen gas, i.e.,

$$U_{strain} = \int_0^\varepsilon \sigma d\varepsilon$$

But, by Hooke's law,

$$\sigma = Y\varepsilon \tag{12.60}$$

where Y is Young's modulus and ε is the strain. Hence,

$$U_{strain} = \int_0^\varepsilon Y\varepsilon \, d\varepsilon = \frac{Y\varepsilon^2}{2}$$

$$= \frac{\sigma^2}{2Y} \tag{12.66}$$

This is the strain energy per unit volume. Hence, for the crack considered, the strain energy is

$$-\frac{\sigma^2}{2Y} V_{crack} \tag{12.66a}$$

If the crack were a sphere (cf. Fig. 12.82), $V_{crack} = (4/3)\pi(l/2)^3$. Let it be taken as $K_1(\pi l^3/8)$, where K_1 is a dimensionless constant close to 1.

But, also, as the crack expands under the influence of the H_2 within it, it absorbs more surface energy.
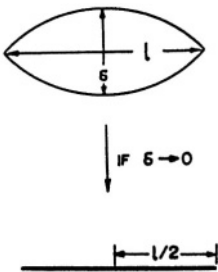
Fig. 12.82. The Griffith crack is a disk-shaped crack. When it is in equilibrium, the strain energy expanding it is equal to the surface energy contracting it. For a quantitative argument, it can be represented by a flat disk of small thickness, δ, and relatively considerable radius $l/2$.

The total surface energy of the crack is

$$\gamma A_{\text{crack}} \tag{12.67}$$

where γ is the surface tension of the metal/hydrogen interface. If the Griffith crack were spherical, A_{crack} would equal $4\pi(l/2)^2$, and if it were a disk, it would be $2\pi(l/2)^2$. Let it be $K_2(\pi l^2/4)$, where $2 < K_2 < 4$. Thus, $\gamma A = K_2\pi(l^2/4)\gamma$. Hence,

$$U_{\text{crack}} = -\frac{\sigma^2}{2Y}K_1\pi\frac{l^3}{8} + K_2\pi\frac{l^2}{4}\gamma \tag{12.68}$$

If the H_2 pressure is large enough, the crack will grow, i.e., l will increase. One can obtain the critical value of σ (the stress in dynes cm^{-2}), and also the H_2 pressure in the crack, by finding what σ value corresponds to:

$$\frac{\partial U}{\partial l} = 0 \tag{12.69}$$

Using (12.65) in (12.63), one finds (assuming $K_2/k_1 \simeq 2$)

$$\sigma_{\text{critical}} = p_{H_2\text{in crack for embrittlement}} > \left[\frac{16}{3}\frac{Y\gamma}{l}\right]^{\frac{1}{2}} \tag{12.70}$$

One may take some likely values of Y, γ, and l to get a feeling of what sort of p_{H_2} values are implied. They are

$$Y = 10^{12} \text{ dyn cm}^{-2}$$

$$\gamma = 10^3 \text{ dyn cm}^{-1}$$

$$l = 10^{-5} \text{ cm}$$

Then,

$$p_{H_2} = \left(\frac{16}{3} \times 10^{20}\right)^{\frac{1}{2}} \text{ dyn cm}^{-2}$$

$$\simeq 3 \times 10^4 \text{ atm} \tag{12.71}$$

Thus, when hydrogen atoms from the metal surface have diffused inside the metal and reached the voids within the metal, they deabsorb from the surface of these H_2molecules. If their pressure inside the voids is high enough (3×10^4 atm, as deduced above), the crack will spread; collapse is on the way.

It is clear that the treatment so far of when the hydrogen pressure inside a void gives rise to the spreading of metals and their breakup is thermodynamic in nature. It implicitly assumes there is an equilibrium between the H_3O ions in solution, the H adsorbed on the electrode surface, the H_2 dissolved in the solution (and that in the gas phase above it), and finally the H dissolved in the metal and the H_2 in the void.

Is this all right? No! what is certain is that there is no equilibrium between the H_3O^+ in solution adjacent to the electrode surface, the H adsorbed on the metal and the H_2 evolved from it into the solution or gas phase.[24] In fact, most of this second volume of our book would not have been written if there were equilibrium in the events at interfaces, as indeed used to be thought by electrochemists before about 1950. So it is necessary to look again at treatments developed on the assumptions of equilibrium in the embrittlement of metals. To be sure, the treatments available in the late 1990s were still not complete because they continued to assume equilibrium between the metal surface, the dissolved H, and the H_2 in the voids to which this H diffuses, but the dangerous simplification about equilibria in the electrode reaction between H_3O^+ in the solution and H_2 in the gas phase has been removed (Bockris and Subramaniam, 1972) and a kinetic treatment, with varying assumptions as to the path and rate-determining step of the hydrogen evolution reaction have been given. The kinetic treatment requires a bit too much space to give it here, but it is possible to encapsulate the results in Table 12.6.

This table is quite enlightening because it differentiates sharply between the various mechanisms of hydrogen evolution and the resulting pressure that would be expected to build up in voids in the metals, assuming (still) that there is equilibrium of H and H_2 *inside* the metal.[25] Thus (recalling that η, the hydrogen overpotential, has a negative sign), it is seen that only four of the six mechanisms listed in the table can give rise to damage to the metal. The two most dangerous surface mechanisms in respect to damage are the fast discharge-slow chemical combination; and the fast discharge-slow electrochemical desorption.

It is possible to make a useful summary equation, for it is seen that five out of the six mechanisms in Table 12.6 are given by an equation of the form:

$$\text{Internal pressure voids} = e^{-x\eta F/RT}, \quad \text{where} \quad 2 > x >> 1/2$$

Accordingly, one sees from Table 12.6 that the thermodynamic treatment gives the highest possible pressure. The general equation for the beginning of electrochemical damage to a metal due to the buildup of molecular H_2 in its voids would be

---

[24]In this discussion, for simplicity it is assumed that the metal undergoing embrittlement is also evolving H_2 and that a net current is passing. However, the source arguments apply if the H is created as the cathodic branch of a corrosion reaction; there is no *net* current.

[25]Of course, inside metals such as Pd, the actual species present is not atomic H, but largely protons. However, as long as the species inside the metal are in equilibrium with molecular hydrogen in the voids, this would not affect the results of Table 12.6.

TABLE 12.6
Summary of the Internal Pressure Calculations as a Function of the Mechanism of the h.e.r.

Mechanism	Fugacity, f_{H_2} (atm)	Remarks
Fast discharge-slow combination	$\exp\left(-\dfrac{2\eta F}{RT}\right)$	Nernst equation valid; could be embrittling
Slow discharge-fast combination	1	Nonembrittling
Slow discharge-fast electrochemical desorption	$\exp\left(\dfrac{2\eta F}{RT}\right)$	Nonembrittling
Fast discharge-slow electrochemical desorption	$\exp\left(-\dfrac{2\eta F}{RT}\right)$	Nernst equation valid; could be embrittling
Coupled discharge-combination	$10^{1.5} \times \exp\left(-\dfrac{\eta F}{2RT}\right)$	$\theta \ll 1$; could be embrittling
Coupled discharge-electrochemical desorption	$\exp\left(-\dfrac{2\eta^* F}{RT}\right)$	Only predictable exactly if η^* is known experimentally[a]

Source: Reprinted from J. O'M. Bockris and S. U. M. Kahn, *Surface Electrochemistry*, 1993, Table 8.11 p. 836, Plenum Publishing.
[a] η^* represents the potential at which fast discharge-electrochemical desorption changes to coupled discharge-electrochemical desorption.

$$e^{-x\eta F/RT} > \left(\frac{16}{3}\frac{\gamma Y}{1}\right)^{1/2} \tag{12.72}$$

This equation tells us that if we wish to know the danger point (i.e., the critical overpotential at which the metal will begin to undergo cracking), we not only have to know γ the surface free energy for the metal; Y, its Young's modulus; and 1, the average "lens length" in the metal's voids before H diffuses into them, but also the applicable mechanism of the hydrogen desorption from the metal's surface, and the rate-determining step in the sequence. This is because the coverage of the metal's surface with H, and the way in which this quantity varies with overpotential, is mechanism dependent.

An interesting point turns up here and that is the role of impurities in solution (particularly organics of various kinds). Such materials adsorb on metal surfaces. Flitt (1981) made a study of these impurities (which could be corrosion inhibiting). Such adsorption changes the mechanism of hydrogen evolution by blocking surface catalytic sites. Thus, one may start out with a certain mechanism at low overpotentials. However, as the hydrogen overpotential is made more negative, the desorption mechanism of H adsorbed to H_2 (gas), may change because of the presumed organic

impurities, the adsorption of which is potential dependent. The point is that such a change in a mechanism for the evolution reaction then changes the pressure inside the voids and alters the potential at which damage will begin according to Eq. (12.72), because the value of H in that equation is mechanism dependent.

How valid is this elementary picture of crack initiation and propagation? It can be tested by measurements of permeation of H inside the metal. One can first make some predictions. Measurements of the rate of travel of atomic hydrogen through a lattice can tell us much about what goes on inside the metal. When the concentration of H in the metal is below a critical limit, the only sinks for the diffusing hydrogen are interstitial traffic positions. However, when the concentration of H inside the lattice exceeds a critical value, then voids present in the structure of the metal near points of triaxial stress will start consuming the diffusing hydrogen. This means that there will be many hydrogen sinks inside the metal and the switching-on of these sinks should reflect itself in the permeation-time behavior. Thus, without them, H is injected from the cathodic parameter regularly from the entry side to the other side of a metal membrane (the "other" side being held at an anodic potential).

Experiment shows that when the hydrogen concentration is enough to cause embrittlement, the permeation current builds up with time and then instead of stabilizing to a steady state as it normally does in the absence of a crack-initiation and propagation process, it *drops down* and only then becomes steady (Fig. 12.83). Thus, the fall in permeation current occurs at the onset of crack propagation and embrittlement.

The permeation-time behavior of a cracked and embrittled membrane differs in another fundamental way from membranes that do not suffer such cracking, i.e., when the H_2 in voids is not high enough in pressure to cause spreading. In the case of the
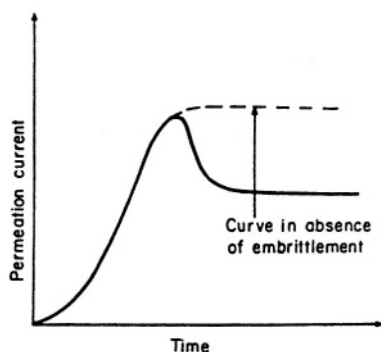
Fig. 12.83. When cracks are initiated and hydrogen starts accumulating inside the cavities, the permeation current falls off.

latter, the permeation-time transients are reversible; after running one transient, the hydrogen inside the metal can be pumped out by dissolution at the exit side (see Fig. 12.70), and a second transient can be shown to reproduce the first. In the case of an embrittled membrane, the second transient is entirely different from the first. This is not due to irreproducibility, as was originally thought. The hydrogen-permeation properties of the membrane have been irreversibly changed by embrittlement (McBreen, 1967). Much of the incoming hydrogen remains trapped inside voids in the metal. If the H_2 in these voids reaches a value greater than the spreading pressure of the metal, the voids will spread and join with other voids. The metal is now broken down inside, and has become embrittled.

Two modes of the internal disintegration of a metal have been described, namely, stress/corrosion cracking and hydrogen embrittlement. What is the difference between the two mechanisms of decay? First, in stress-corrosion cracking, stress (either applied from outside or residual from manufacture of the component) is a necessary but not sufficient condition, whereas in hydrogen embrittlement, the local stress is caused by H_2 at high pressure in voids. The permeation of hydrogen into the metal is a necessary condition for hydrogen embrittlement. In stress-corrosion cracking, it is the surface crack that is a necessary condition. It may be the result of mechanical stresses or it may arise from preferential hydrogen entry in regions of stress. Thus, hydrogen has a sporadic role in stress-corrosion cracking.

Finally, there is the question of the electrochemical basis of the two modes of internal decay of metals. The propagation of a stress-corrosion crack in metals is *sustained* by the electrochemical metal-dissolution reaction, although much of the net motion is due to the exit of the metal through regions in which an extremely strong concentration of H has been has been caused by local stress. On the other hand, one can conceive of hydrogen embrittlement even though the permeating hydrogen does not arise from an electrochemical hydrogen-evolution reaction at the metal surface. For example, metals embrittle in the presence of hot, dry hydrogen. However, under practical conditions, hydrogen is usually introduced into the metal by the electrodic hydrogen-evolution reaction that unintentionally occurs on the exposed external surface as part of a corrosion couple.

12.7. WHAT IS THE DIRECT EXPERIMENTAL EVIDENCE FOR VERY HIGH PRESSURES IN VOIDS IN METALS?

12.7.1. Introduction

The account of damage to metals by renegade protons or H atoms that weaken bonds within the metal is acceptable enough, but that part of the picture which goes a bit further and puts the H atoms up against the edge of a void, finally recombining on its surface and forming molecular hydrogen at high pressure inside the void, has strained the confidence of some scientists in the formulas. There is a good reason for

this. The pressures one calculates when one applies some of the formulas of Table 12.6 are huge indeed. Consider, for example, the coupled discharge-combination mechanism for the hydrogen-evolution reaction. According to Table 12.6, the equation for the H_2 pressure inside voids is

$$10^{1.5} e^{-\eta F / 2RT}$$

Then, with $\eta = -0.5$, the pressure inside a void would be about 10^4 atm.

It is no use trying to apply the ideal equation of state for gases, $pV = uRT$, to such high pressures. There are empirical equations (the Beatty–Bridgemann equation is one) that allow one to find the number of moles of H_2 in a void of known volume at such pressures where the moles present may be as low as 10^{-12} at the beginning before the crack has spread.

In choosing the coupled discharge-combination mechanism for the mechanism of hydrogen evolution (Jackson,), the *lowest* pressures corresponding to a given over-potential are calculated. Other mechanisms for the hydrogen evolution reaction indicate that still greater hydrogen pressures than the 10^4 atm indicated should be effective in voids.[26] The predictions of such high pressures in voids by these electrochemical equations is very challenging indeed and it is of interest to consider to what degree there is experimental evidence for them.

12.7.2. A Partial Experimental Verification of High Pressures in Metal Voids

High pressures can be assayed in their buildup phase by using a macroversion of such voids (Fig. 12.84). The apparatus (Minevski and Lin, 1998) shows an electrochemical cell in which the cathode is a cylinder of palladium (chosen for its large permeability coefficient for the throughput of H) that contains a void space into which H_2 diffuses from the deposition of H on the surface of the cylinder. This space is connected to a pressure transducer, which works on the principle of a piezoresistor, i.e., it contains a substance that produces an electrical potential upon being compressed. The rise of pressure inside the void space is shown in Fig. 12.85.

The pressure in the macro void space depends significantly upon the overpotential, but rises only slowly to the final value expected after equilibrium between the surface and molecular hydrogen has been established. Thus, the principle of an electrode generating H_2 at 1 atm on the outside and providing pressures in "voids" of several hundred atmospheres is proven. However, the existence of void pressures of *thousands* of atmospheres (as predicted by the equations in Table 12.6) is not proven

---

[26]The super-high pressures that arise by application of the equations in Table 12.6 are fugacities. The difference between the fugacity of a gas and its pressure occurs because of the deviation from the ideal ($PV=nRT$) behavior of gases, which happens increasingly when the gas pressure exceeds ~100 atm. Thus, a super-high fugacity (10^{10} atm, for example) may be equivalent to a pressure several orders of magnitude smaller.
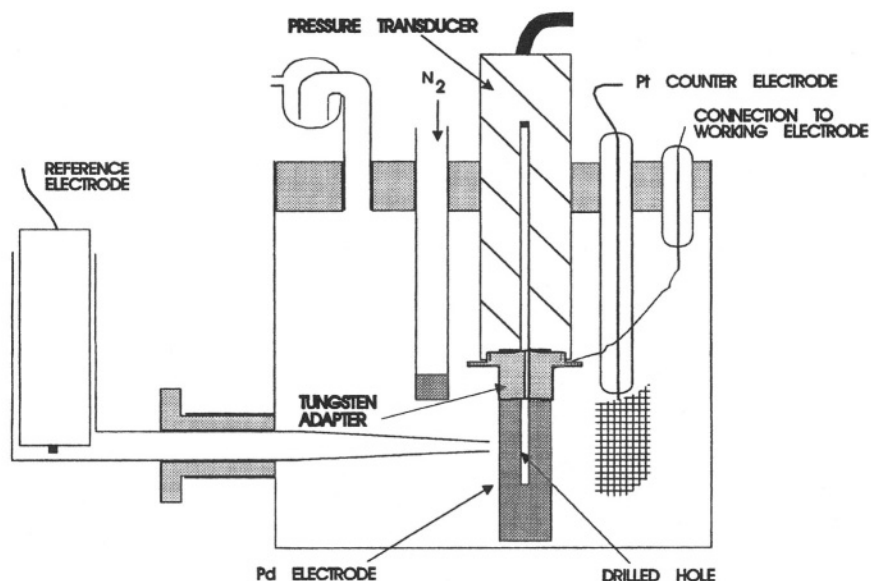
Fig. 12.84. Cell for high-pressure measurements. The "drilled hole" constitutes the void space into which the H_2 diffuses. The H_2 gas in the void space arises from H atoms deposited on the outside of the Pd cylinder, which functions as an electrode when the cell above is supplied with solution and the current is turned on. (Reprinted from Z. Minevski, dissertation, Texas A&M University, 1995.)
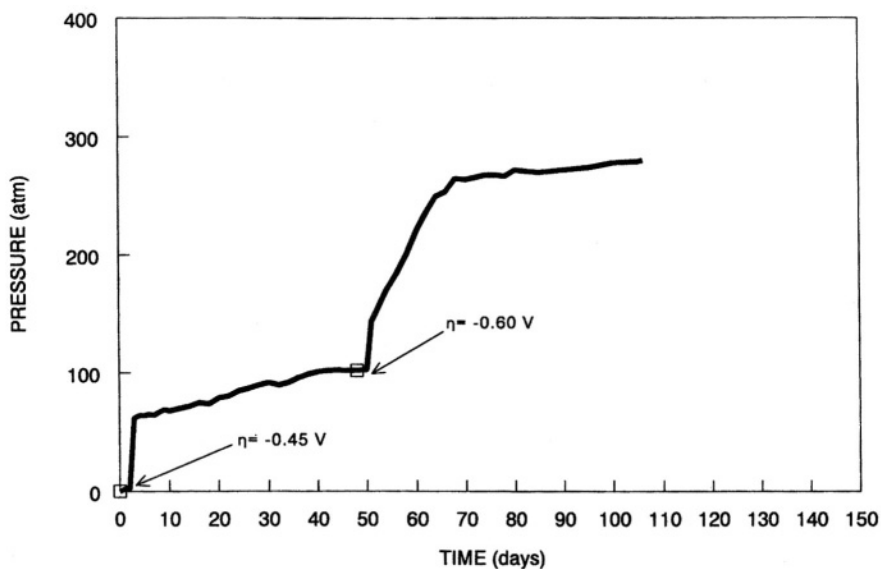
Fig. 12.85. Pressure transient obtained during the sequence of increasing overpotentials: $\eta = -0.45$ V and -0.60 V. (Reprinted from Z. Minevski, dissertation, Texas A&M University, 1995.)

by this direct approach because of the long time needed to reach equilibrium between the outside surface of the electrode in contact with the solution through the relatively thick electrode walls (Minevski, 1998).

12.7.3. Indirect Measurement of High Pressures in Voids

It is possible to artificially produce a number of voids in Pd, each ~50 μm in diameter. A layer of Pd can be deposited over the voids that remained empty. This can be done by filling the cavities with a high molecular weight (low vapor pressure) hydrocarbon (octacosane), evaporating Pd over the cavity and then evaporating out the octacosane through a pinhole in the Pd, this to be filled later by electrodeposition of a thick layer of Pd. Enough hydrogen can then be introduced on the cathodic side (see Fig. 12.86), to fill the artificial voids in the Pd electrode with H_2 at the pressures implied in Table 12.6. The electrode is then used at a bielectrode, the "back" side being polarized anodically to remove the hydrogen in the artificial voids in about a day.
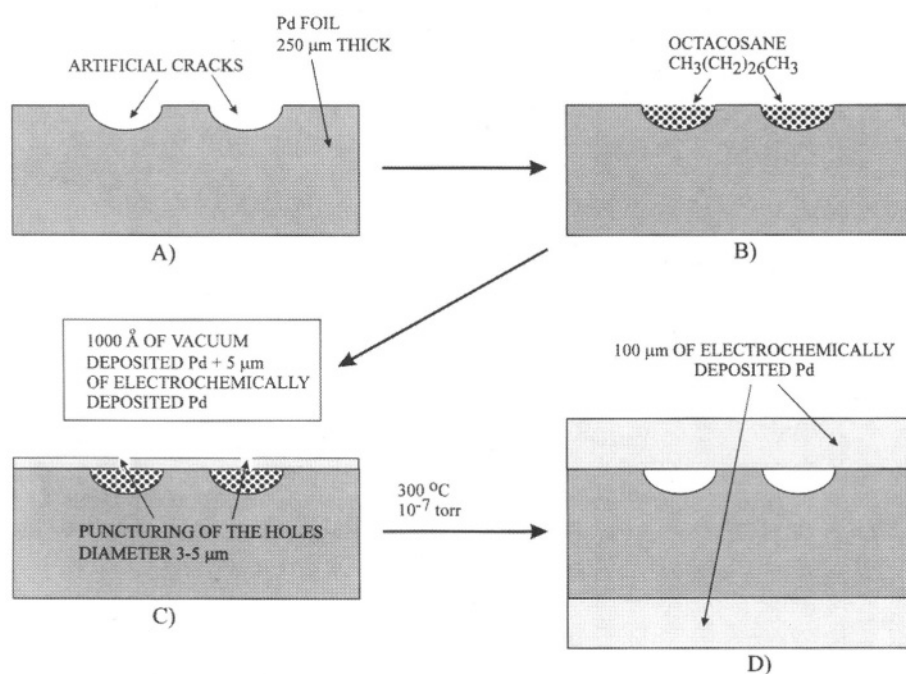
Fig. 12.86. Preparation of Pd membrane with artificial voids. (A) indented artificial voids, (B) voids filled with octacosane, (C) microcavities in 5-μm-thick electrodeposited Pd, (D) electrode covered with 100-μm-thick electrodeposited Pd layer. (Reprinted from Z. Minevski, dissertation, Texas A&M University, 1995.)

Figure 12.87 shows the anodic decay current for the remaining H_2 from the voids as a function of time. It is seen that "bumps" occur on such curves, and from an examination of the behavior of this anodic current-time decay curve, it is possible to obtain the calculated pressure in the voids. The result of this indirect approach was 3600 atm for an overpotential of –0.6 V in 1 M NaOH (Minevski and Lin, 1998).

Such an experiment gives some support for the very high pressure in voids which is the basis of some of the hydrogen embrittlement models discussed here. However, the success of this experiment should not be taken as indicating that damage inside metal is entirely and only due to the mechanical force of high pressure in voids. For example, the H adsorbed on the surface of grains in the crystals of Pd will reduce the adhesion between grains and in this way contribute to the breakdown of the lattice (Petch and Stables, 1952; Oriani, 1989).
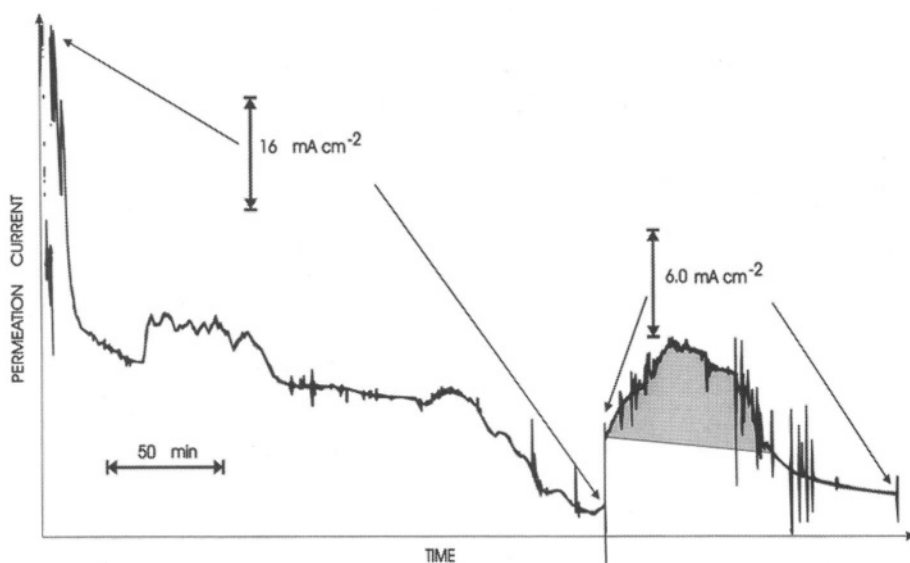
Fig. 12.87. A decay transient from the work of Minevski et al. The current observed is an anodic current arising from the dissolution of hydrogen from palladium after an earlier saturation of the palladium during H_2 evolution. Note the unusual "bump" on the current–time line. Analysis of this area of the transient gave rise to a calculation of the equivalent amount of hydrogen it represents. Thus, knowing the voltage from which it came and using appropriate equations of state, it is possible to calculate the pressure in the voids from which it originated. (Reprinted from Z. Minevski, dissertation, Texas A&M University, 1995.)

12.7.4. Damage Caused Internally in Metals by the Presence of H (and H$_2$) at Varying Overpotentials

It has been known for about a century that the evolution of hydrogen on the surface of metals may lead to internal damage. The dependence of the extent and type of this damage upon the overpotential at which the hydrogen evolution occurs, and the temperature and the depth of penetration from the surface of the damage at a given time, are of substantial importance, although they have been little examined from a fundamental viewpoint. For example, hydrogen evolution at a fixed overpotential can be allowed to occur for varying periods, potentials, and temperatures, and specimens of the metal thus exposed can then be etched to various depths and the internal structure examined (e.g., by polarization interference microscopy and scanning electron microscopy) (Minevski, 1994).

Palladium is a metal of particular interest in this respect because it has an unusually high solubility for H and because it is the principal metal in which chemically assisted nuclear effects have been reported (Fleischmann and Pons, 1989). For the evolution on Pd of H$_2$ from LiOH and LiOD-containing solutions,
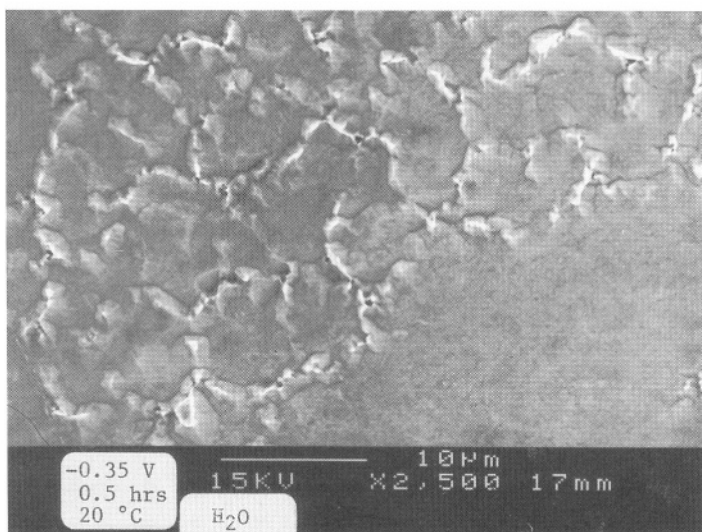
Fig. 12.88. SEM micrograph of Pd electrode electrolyzed for 30 min at overpotential of η = –0.35 V and temperature of T = 20 °C. Magnification ×2500. (Reprinted from Z. Minevski, dissertation, Texas A&M University, 1995.)

no internal damage can be observed after up to 6 weeks of electrolysis as long as the overpotential is maintained less negative than –0.25 V. Damage to the metal structure is observed at greater overpotentials, but after 3 weeks of continuous electrolysis at the very high overpotential of 1 V (and at temperatures up to 80 °C) damage was not observed at a greater depth than 10 µm. There were several characteristics of the damage observed:

1. No damage was visible on the electrode surface (as indicated by SEM up to × 10,000) after 3 weeks at $\eta = 1.0$ and T = 300 °K.
2. A pattern of hexagonally positioned "pinholes" emerged along grain boundaries in the metal (Fig. 12.88). The extent of these formations increased linearly with time and somewhat exponentially with overpotential and temperature.
3. The most common type of damage found was that of irregularly shaped caverns (Fig. 12.89) (up to 1 µm in diameter). However, these formations began in island structures that eventually spread (over 2–3 weeks for $\eta = -1.0$ and T = 300 °K) over most of the electrode to 1 µm in depth.
4. After 6 weeks of electrolysis at a –1.0 V overpotential, pitting of the metal could finally be observed on the surface.
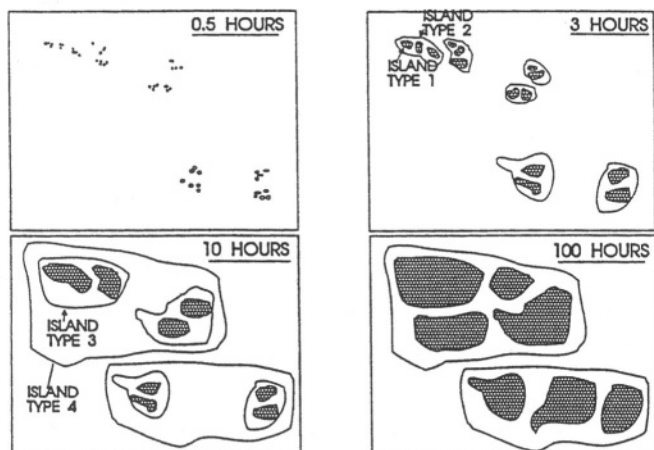
Fig. 12.89. Schematic representation of subsurface changes of Pd electrodes electrolyzed at overpotential of $\eta = -1.00$ V and temperature of T = 20 °C for 30 min, 3 hr, 10 hr, and 100 hr of electrolysis.

Further Reading

Because the study of the phenomena of the interaction of H with metals began at an early date, there is an extremely large literature (more than 2000 papers) and our choices have had to be particularly selective.

Historical and Seminal

1. T. Grahame, *Phil. Trans. Roy. Soc. London* **156**: 415 (1860). *Possibly* prior to Cailleset's discovery (1864), Grahame discovered that metal absorbs H (which he called hydrogenium).
2. L. Caslletet, *Compt. Rend.* **56**: 327 (1864). Reported disappearance of H inside iron during pickling.
3. A. Sieverts, *Z. Physikal. Chem.* **60**: 169 (1907). Establishment of the law pertaining to the amount of dissolved H and the H_2 external pressure.
4. A. Griffith, *Phil. Trans.* **A221**: 102 (1920). The lenslike shape of voids in metals.
5. C. A. Zappfe and C. E. Sims, *Trans. ASME* **255**: 145 (1941). The first suggestion of high pressure in voids in metals as a mechanism of embrittlement.
6. A. N. Frumkin and N. Aladjalowa, *Acta Physicochem. USSR* **19**: 1 (1944). First measurements on anodic H on positive side of bielectrode.
7. N. J. Petch, *Phil. Mag.* **3**: 1089 (1958). Decay of properties due to H adsorption at grain boundaries reduces surface bonding between grains.
8. Z. Szarklarska-Schmiolowski and M. Smialowski, *Bull. Acad. Pol. Sci. Ser. Chim.* **6**: 247 (1958). H solubility in iron.
9. A. R. Troiano, *Trans. ASM* **54**: 52 (1960). First suggestion of the congregation of H at points of triaxial (i.e., high) stress points.
10. M. A. V. Devanathan and Z. Stachurski, *Proc. Roy. Soc. London* **270**A: 96 (1962). Theory of the cell for the electrochemical determination of H damage in metals.
11. A. S. Tetelmann and W. D. Robertson, *Acta Met.* **11**: 415 (1963). Pressure theory, quantitative.
12. W. Beck, J. O'M. Bockris, J. McBreen, and L. Nanis, *Proc. Roy. Soc. London* **A290**: 220 (1966). Partial molar volume of H in iron. Relation of solubility to local stress. Permeation as an arbiter of damage.
13. R. G. Raicheff, A. Damjanovic, and J. O'M. Bockris, *J. Chem. Phys.* **49**: 926 (1968). The effect of stressing metals upon the rate of appearance of slip planes of different indices.
14. A. R. Despic, R. C. Raicheff, and J. O'M. Bockris, *J. Chem. Phys.* **49**: 926 (1968). Effect of stress and yielding in metals upon the anodic current density.
15. J. O'M. Bockris and P. K. Subramanyam, *J. Electrochem. Soc.* **118**: 114 (1971). H_2 traps the pressure produced.
16. J. O'M. Bockris and P. K. Subramanyam, *Electrochim. Acta* **16**: 2169 (1971). Internal pressure as a function of overpotential for various kinetic mechanisms of the surface desorption of H.

Modern

1. H. J. Flitt and J. O'M. Bockris, *Int. J. Hydrogen Energy* **7**: 411 (1982). Effect of organic inhibitors on the ingress of H into metals.

2. H. J. Flitt and J. O'M. Bockris, *Int. J. Hydrogen Energy* **8**: 39 (1983). A laser-based technique for measuring H in local areas of metals.

3. T. B. Flanagan, *Proc. Electrochem. Soc.* **94-21**: 17 (1995). Cathodic absorption of H; review by a principal contributor to the field.

4. G. Jerkiewicz, J. Borodzinski, W. Chrzanowski, and B. E. Conway, *Proc. Electrochem. Soc.* **94-21**: 44 (1995). Factors involving blocking of H absorption.

5. M. Enyo, *Proc. Electrochem. Soc.* **94-21**: 75 (1995). H pressure in cathodes.

6. O. Yamazardi, H. Yoshitaka, N. Kamiya, and K. Ohta, *Proc. Electrochem. Soc.* **94-21**: 92 (1995). H absorption as a function of Li inclusions in Pd.

7. F. R. Durand, J. C. Chen, J. P. Dicard, and C. Montella, *Proc. Electrochem. Soc.* **94-21**: 207 (1995). Impedance study of H absorption.

8. E. Protopopoff and P. Marcus, *Proc. Electrochem. Soc.* **94-21**:3 74 (1995). Site blocking of H entry.

9. L. J. Gao and B. E. Conway, *Proc. Electrochem. Soc.* **94-21**: 388 (1995). Poisoning of H entry into metals.

10. J. O'M. Bockris, Z. Minevski, and G. H. Lin, *Proc. Electrochem. Soc.* **94-21**: 410 (1995). The experimental establishment of 3000 atm pressure in voids in Pd.

12.8. FATIGUE

Like the invisible spreading of internal cracks that can bring down bridges and sink ships, fatigue is an insidious danger to the integrity of certain kinds of metallic structures. It is a general phenomenon in which loss of strength comes with repeated cyclical stressing. It is greatly worsened by the contact of the metal with an electrolytic solution, namely, sea water.

The most famous (and most dramatic) example of the effect of fatigue failure is that of the Comet series of aircraft, the first commercial jets, which were operated by British Airways. A total of four such aircraft fell apart in flight before the design was withdrawn from service. The fuselage of a remaining aircraft of the series was subsequently enclosed in a tank and rhythmically stressed until a crack was observed to develop. Under the stressed conditions of flight, such cracks would spread rapidly and the fuselage would burst open. Landing and flight are equivalent to subjecting an aircraft fuselage to rhythmic pressure as a consequence of the reduction of atmospheric pressure at high altitudes.[27]

---

[27]A further example of the dramatic effects of stress is given by a VC-10 aircraft of the same company. While flying over Mount Fuji, Japan, the tail assembly of one such plane became detached from the rest of the jet. Turbulence was the stated cause. However, H related fatigue was probably a forerunner of the disaster.

The effect of an electrolytic solution on fatigue failures is exemplified by the behavior of ships' propellers. If they are out of alignment, the stress that this brings may occur in a region of the propeller shaft that is dry. Failure may then occur only after months. However, if a leak allows the stressed part to become wet, failure occurs in days.

In a general way, fatigue failure may be interpreted in terms of the enhancement of the diffusion of vacancies to grain boundaries. If sufficient vacancies arrive there, the cohesion between grains is reduced and the metal's strength lost.

Insofar as electrolytic effects are concerned, cyclical stressing would cause high index slip steps to emerge preferentially at the metal surface (see Damjanovic and Raicev, 1966). Adsorption of ions on the freshly exposed surface would prevent the slip steps from retreating into the lattice on the reverse cycle. However, high index planes have abnormally high i_0's for metal dissolution. Corrosion is therefore enhanced in cyclically stressed areas. The effect of sea water would follow as a consequence of Cl^- effects in breaking down the protective passive layers.

Enhanced dissolution by mechanisms such as the above introduce vacancies and divacancies on the surface which diffuse back into the metal to grain boundaries, where they cause enhanced creep and plastic deformation. This pressure of these would eventually lead to failure (Uhlig and Revie, 1985). An H embrittlement effect may add to this mechanism (Thomas and Wei, 1992).

12.9. THE PREFERENTIAL FLOTATION OF MINERALS: AN APPLICATION OF THE MIXED POTENTIAL CONCEPT

12.9.1. Description

Minerals do not lie neatly each in its own bundle in the ground. When a valuable sulfide or oxide is discovered, the desired material is mixed with earth and rock and other minerals. How is the desired mineral in a mixture to be separated from its undesirable companions?

The solution found to this dilemma has been a process called *froth flotation,* first developed in Australia in the 1900s. The ore, which contains several different minerals, is crashed (thus liberating grains of the minerals) and made into a pulp with water. A specific organic species that has been found to adsorb selectively only on the surface of the grains of the valuable mineral species is then added. This compound renders the surface of the desired ore hydrophobic. Air is bubbled through the pulped ore mixture containing the adsorbed organic. Air bubbles attach themselves selectively to the surface of the one ore made hydrophobic and lift the particles containing it to the surface of the bath; there they make a froth layer that contains the desired mineral on the surface of the liquid in the flotation cell. The desired separation is thus effected. Typically, the selected ores may be galena (lead sulfide) or chalcopyrite (cuprous

sulfide), and the organic (called the collector) a xanthate, RO–C–SNa, where R is an alkyl group (Keller and Lewis, 1925).

The creation of a nonwettable surface depends on that surface having a solid/solution surface tension that causes the contact angle between it and the solution to be changed to that corresponding to the nonwettability of the surface. Looking back to the description of electrocapillary curves (Section 6.5.2), it can be seen that the surface tension depends on the potential and the substance adsorbed. Thus, if the surface of, e.g., galena is at first hydrophilic, and then after addition of the collector it becomes hydrophobic, it is the adsorption of the organic xanthate collector that must have caused the surface tension change that has brought about the hydrophobicity, and thus enabled the air bubbles to attach themselves to the ore, etc.

What, then, is the part of the process that justifies the heading of this section and brings the material into a chapter on materials science in electrochemistry? It is the process by which xanthates adsorb. It has been established (Nixon, 1957) that the formation of the monolayer of an organic substance is not a physical but a chemical, indeed an electrochemical, process. The xanthate undergoes an anodic oxidation:

Xanthate ion in solution,   Dixanthate
adsorbed on mineral
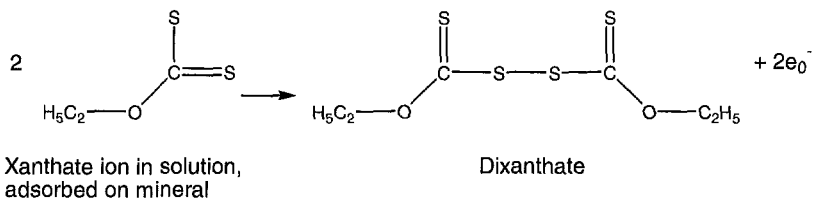
However, for the anodic reaction to occur, there must be a corresponding cathodic reaction—hence the reference to corrosion and the Wagner–Traud hypothesis. What is the counter-reaction that takes up the electrons rejected to the underlying semiconductor, thus setting up a mixed potential?

The answer was given in a hypothesis formulated by Salamy and Nixon (1954). They suggested that the anodic reaction (12.69) was accompanied by a cathodic partner reaction, the reduction of O_2:

$$1/2 \; O_2 + 2H^+ + 2e \rightarrow H_2O \tag{12.73}$$

Thus, were the xanthate ion itself to adsorb and retain its charge, lateral repulsion would make it impossible for the surface coverage on the mineral to be a high one, and the desired hydrophobicity of the surface would not be achieved. In the electrochemical mechanism described by Salami and Nixon, the adsorption can become a charge-transfer reaction, continuing by the participation of oxygen until the surface is fully covered with dixanthate (and hence wettable). The mechanism is thus an electrochemical oxidation.
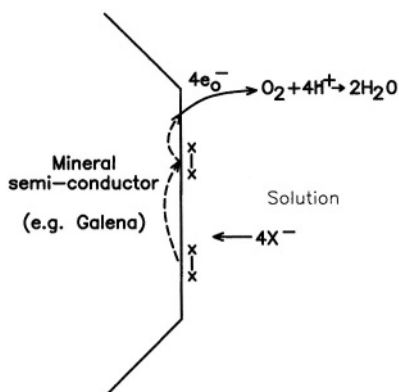
Fig. 12.90. The adsorption of the xanthate occurs as a charge-transfer reaction made possible by the reduction of oxygen and leading to the adsorption of dixanthate, which makes the surface hydrophoric and hence attractive to bubble adhesion.

Supporting evidence for this seminal concept was provided by Tolun and Kitchener (1964) when they found that the potential attained by galena in the presence of xanthate and oxygen was between that at which O_2 was reduced on galena in the absence of xanthate and that at which xanthate was oxidized on galena in the absence of O_2. Thus the mechanism of formation of the adsorbed hydrophobic, floatable layer was a mixed potential-type process, the same as that originally suggested by Wagner and Traud[28] for corrosion (see Fig. 12.90).

The mixed-potential theory of mineral flotation was established quantitatively by direct examination (Pillai and Bockris, 1984). O_2 was reduced on galena, establishing the Tafel line for this reaction (no xanthate present), and then the oxidation of xanthate was examined, establishing *its* Tafel line in the absence of oxygen. Putting these two lines together should give their intersection at the mixed potential i.e., the open-circuit potential at which galena (pyrites, etc.) floats in the presence of xanthate. At first, when the two Tafels were compared, there was only qualitative agreement, but when the

---

[28]Indeed, the mathematical form of the mixed-potential concept (Bockris, 1954) has been applied to a number of chemical processes which, it has been shown, in fact, consist of two partnered surface electrochemical processes (Spiro, 1984). Thus, energy conversion processes at the surface of mitochondrial cells may involve the electrochemical oxidation of glucose as the anodic reaction and the electrochemical reduction of oxygen as the cathodic (Gutmann, 1985).
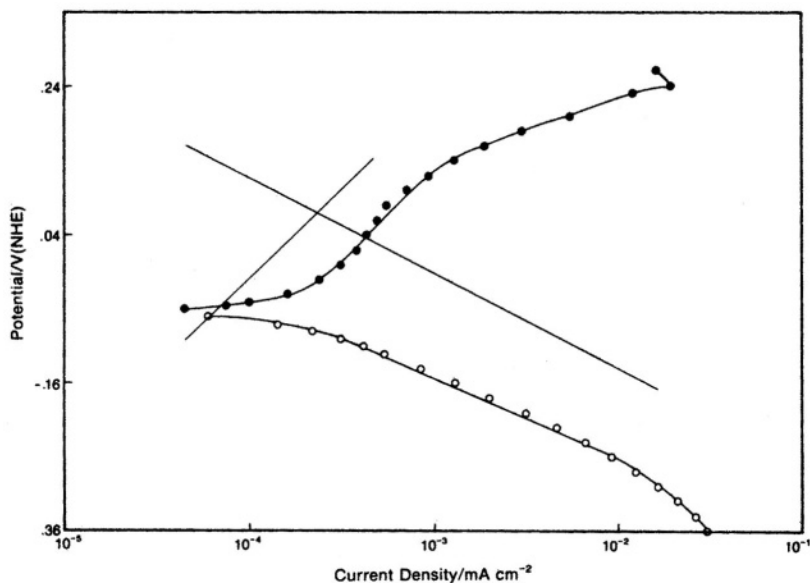
Fig. 12.91. *E* log *i* plots at a PbS electrode in borate buffer pH 9.1. -o-o-; O_2 reduction ($p_{O_2} = 1$) in the presence of 10^{-2} *M* potassium ethyl xanthate (KEX). -•-•-: KEX oxidation ($C_{KEX} = 10^{-2}$ *M*) in the presence of O_2 of $p_{O_2} = 1$. • • • •: Measurements in the absence of other components. (Reprinted from K. C. Pillai and J. O'M. Bockris, *J. Electrochem. Soc.* **131**:572, 1984, Fig. 8. Reproduced with permission of the Electrochemical Society, Inc.)

effect of product of the oxidation of xanthate (which rapidly covers the electrode) was taken into account, a quantitative match of the results (Fig. 12.91) with theory could be obtained. Among all the applications of surface chemistry, separation of ores by froth flotation has the greatest financial value (Woods, 1996).

12.10. AT THE CUTTING EDGE OF CORROSION RESEARCH: THE USE OF STM AND ATM

12.10.1. Application

Until the beginning of the 1990s (see Sonnenfeld and Hansma, 1986), the events at surfaces during corrosion had to be deduced by interpreting the results of many kinds of electrochemical measurements, including some spectroscopic ones. Although much has been learned by these methods (as illustrated in the preceding sections of this chapter), the advent of scanning tunneling microscopy and atomic force micros-

copy (Section 7.5.18) has provided tools that approach those that would be available if the researcher to had an angstrom-scale microscope. By the century's end, we were still in the early stages of learning what these revolutionary tools can tell us about corrosion. It is too early to give a complete description of what that is. Some examples must suffice.

Figure 12.92 shows Fe in air. The scale is 72×72Å per cm in the x and y directions and 40 Å per cm in the vertical direction. Domelike formations and a layerlike growth (of air formed oxide) are seen. After immersion for 3 min in a pH 8.4 buffer solution and being held at -0.75 NHS (at which it is known the reduction of iron oxide occurs), it can be seen (Fig. 12.93) that the domelike structures are lessened, although on the right, the surface is quite rough, with 10 Å promontories. After 8 min, the surface has become much smoother, i.e., the oxide has been largely reduced and the bare surface can be seen (Fig. 12.94). When, the potential is returned to -0.45 vs. NHS (a region for oxide film growth), one does indeed see a return of the hemispherical oxide humps (10 Å high). Further cycling of the electrode between potential regions in the borate buffer in which oxide films grow (-0.455) and are reduced (-0.75) suggests that oxide growth begins in patches with dimensions in the tens of angstroms and that these eventually fuse together to form a smooth layer of oxide (Bhardwaj and Gonzalez-Martin, 1991).
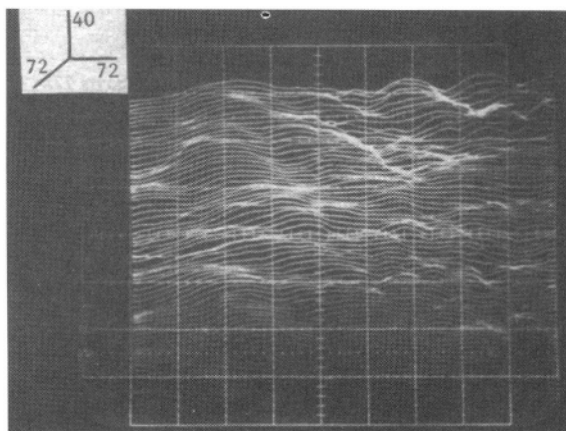
Fig. 12.92. STM image of iron in air. $V_{bias} = -0.298$ V. (Reprinted from R. C. Bhardwaj, A. Gonzalez-Martin, and J. O'M. Bockris, "*In Situ* Scanning Tunneling Microscopy Studies on Passivation of Polycrystallin Iron in Borate Buffer," *J. Electrochem. Soc.* **307**: 1902, 1991. Reproduced with permission of The Electrochemical Society, Inc.)
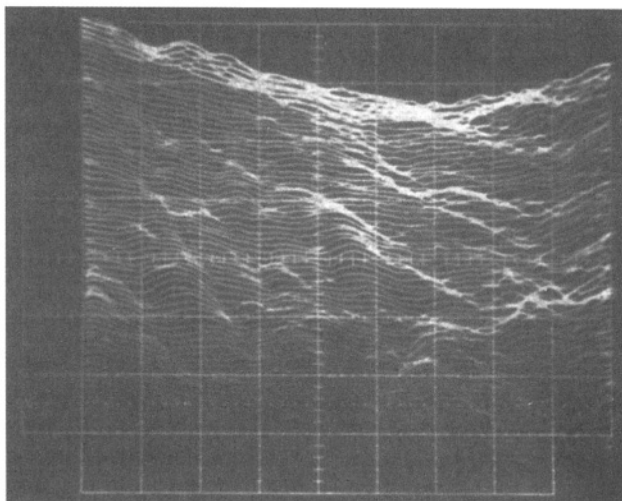
Fig. 12.93. STM image of iron in solution (borate buffer, pH = 8.4) after 3 min at $V_{Fe} = -0.75$ V. (Reprinted from R. C. Bhardwaj, A. Gonzalez-Martin, and J. O'M. Bockris, "*In Situ Scanning Tunneling Microscopy Studies on Passivation of Polycrystallin Iron in Borate Buffer*," *J. Electrochem. Soc.* **138**: 1903, 1991. Reproduced with permission of The Electrochemical Society, Inc.)

The other example concerns alloys of Al with Cu and Zu. Such alloys are widely used for aerospace purposes, so that their corrosion, particularly as it relates to strength of the metal under stress, is of special concern. Environments that aircraft meet on the ground in areas (e.g., Los Angeles) undergoing a smog attack are equivalent to contact of the exterior of the vehicle with moisture films of pH as low as 3!

In this example (Section 7.5.18) AFM rather than STM was used. Whereas STM depends on electron transfer between the specimen and the tip of the probe, AFM depends only on mechanical forces and is independent of the conductance of the specimen, and this may be an advantage for alloys of Al with oxides that show poor conductance. In other experiments AFM images of an Al-Cu alloy immersed in 1 M HCl were recorded. After 24 hr new pits were formed and the ones formed earlier have grown. After 6 hr the sample was severely damaged and the surface is very rough.

Pits are found to develop particularly in the region of grain boundaries which themselves are readily open to attack ("intergranular corrosion"). Thus, it is the Cu-depleted regions that are the major cause of the intergranular damage of Al-Cu alloys. The Al itself remains passive for a certain time and potential, but even when
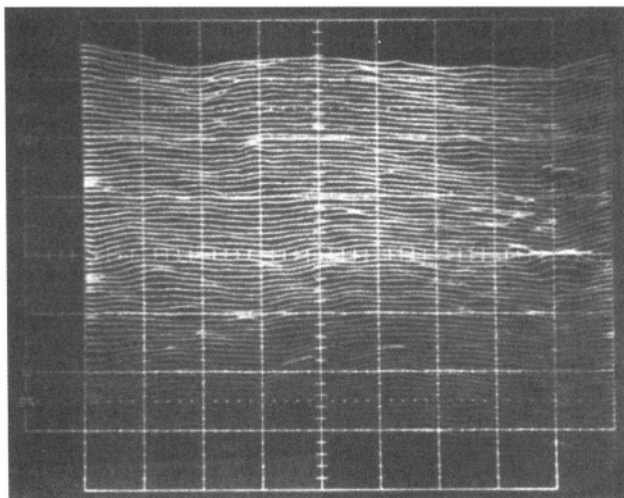
Fig. 12.94. STM image of iron in solution after 8 min at $V_{Fe} =$ –0.75 V. (Reprinted from R. C. Bhardwaj, A. Gonzalez-Martin, and J. O'M. Bockris, "*In Situ* Scanning Tunneling Microscopy Studies on Passivation of Polycrystallin Iron in Borate Buffer," *J. Electrochem. Soc.* **138**:1902, 1991. Reproduced with permission from The Electrochemical Society, Inc.)

the passivity eventually fails to protect it, dissolution along grain boundaries is faster than that of the surface of the grains away from the boundaries.

These examples (Farrington, 1996) illustrate the detail possible with these tools and the extremely heterogeneous nature of real corrosion in contrast to the image used in the Wagner–Traud model of corrosion, i.e., a uniform plane.

12.11. A LASER-BASED TECHNIQUE FOR THE QUANTITATIVE MEASUREMENT OF H IN LOCAL AREAS

12.11.1. Description

One of the lessons learned as corrosion research entered the second half of the twentieth century concerned the importance of making a map showing the heterogeneous zones of most corroding surfaces. Subramaniam (1983) showed that the solubility of H in a metal increases exponentially, depending on the local tensile stress. Stresses are particularly severe near dislocation buildups in metals. The need for a method to examine these theoretical deductions experimentally, and indeed to examine the sites of H in metals, was the origin of the technique described here.

The central idea in the laser-based method for detection of small quantities of H in restricted areas is that an intense laser "strike" can form a "pothole" in a metal. The metal of the pothole and the H it contains are vaporized. Both the H and some of the metallic atoms are withdrawn by a vacuum that pulls gaseous constituents out of the system. On the way, however, these constituents are made to fly through the space between the electrodes of a quadrupole mass spectrometer set to measure mass 2. Knowing the H_2 produced from one laser strike, and the dimensions of the pothole, the H concentration in the metal—and within any restricted area to which the laser can be directed—can be measured.

The laser is used in two ways. In the first strike, it is defocused and its beam is used simply to clear adsorbed impurities, films, etc., from the area, a small portion of which (say, a micron in dimension) is to be the object of a pothole excavation. In the second strike, the laser is intensely focused after passing through a lens system, and the intensity of its strike vaporizes metal to form a hole. The size of the potholes is as little as 0.1 μm in diameter and about 1 μm in depth. A hard vacuum (10^{-10} to 10^{-9} mm Hg) is necessary to give the quadrupole mass spectrometer the required sensitivity. A diagram of the technique is shown in Fig. 12.95.

To test the validity of the present laser-based results, the content of H absorbed into the steel by electrolysis at a series of overpotentials was extrapolated so that the
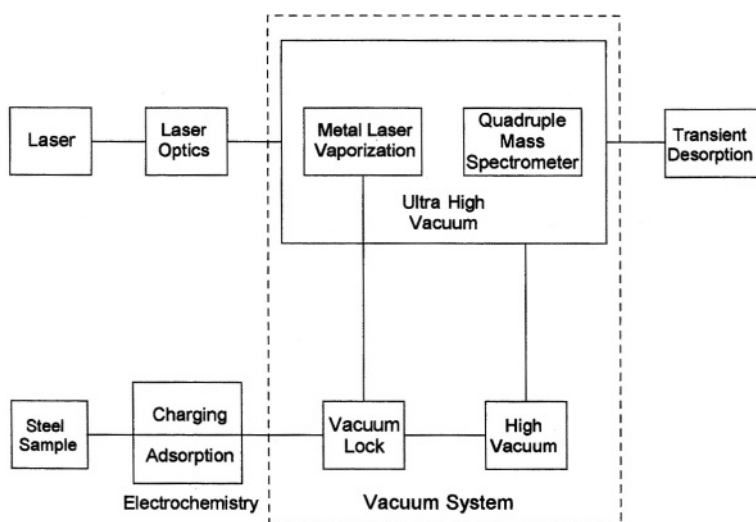
Fig. 12.95. Block diagram of experimental scheme. (Reprinted from H. J. Flitt, J. Pezy, and J. O'M Bockris, "A Laser-Based Technique for the Measurement of Hydrogen at Local Areas in Metals" *Inter. J. Hydrogen Energy* **8**: 40, 1983, with permission from the International Association for Hydrogen Energy.)

value corresponding to zero overpotential (i.e., equilibrium) and this value could be compared with that obtained from gas phase H_2 dissolution in the steel. The consistency between the laser-determined H concentration and that obtained independently by other methods is shown in Fig. 12.96.

The limit of the method is given by the quality of the optics. A sensitivity to areas on the order of $0.1 \mu m$ in size has been achieved. A comparison of this laser-based technique with that of tritium radiography has yet to be made.

12.12. OTHER METHODS OF EXAMINING LOCAL CORROSION

12.12.1. Description

There are several other methods [apart from microellipsometry (Schulze, 1976)] of determining the rate and location of the decay of materials. Two of them are mentioned here; the rest are left to extra reading. The two seem to be related, but their only connection is that each can be used to examine corrosion.

Both these methods are described with the word "noise," but in the first the "noise" has nothing to do with sound; it refers to noise in the sense one meets the term in electronics. In electrochemistry, it refers to the random variation of the electrode potential, which has an order of magnitude of 10^{-4} V and a bandwidth of ~1 Hz.
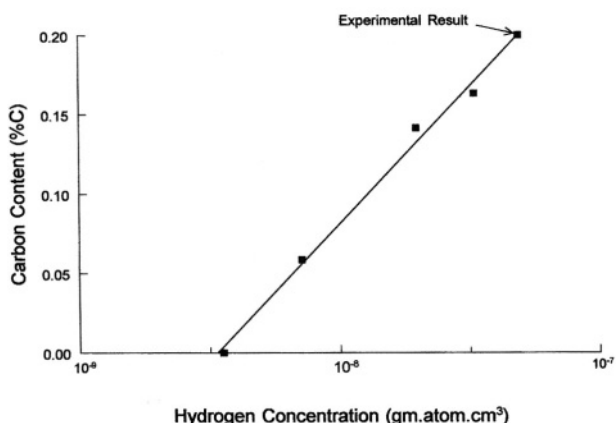
Fig. 12.96. Concentration of hydrogen in steels of various carbon contents, compared with the result from present analysis extrapolated to charging or zero overpotential (Reprinted from H. J. Flitt, J. Pezy, and J. O'M Bockris, "A Laser-Based Technique for the Measurement of Hydrogen at Local Areas in Metals" *Inter. J. Hydrogen Energy* **8**:47, 1983, with permission from the International Association for Hydrogen Energy.)

Measurement of the potential noise at an electrode can lead (though there are not a few assumptions) to the determination of the current passing across the electrode/solution interface, and hence, in a corroding electrode, to the corrosion current. It turns out that the corrosion current density is proportional to the reciprocal of the mean square of the noise.

Measurements of noise can be applied throughout electrochemistry (Tyagai, 1967), but the correlation between the quantity thus determined and the same quantity determined by a more conventional route leaves much to be desired. So why bring up the method here? It has one big advantage. Like the decay method of following transients in electrode measurements (Section 8.4.1), it is independent of IR drop. Thus, it might be the only method possible in a system with poor conductance (Langyal, 1996).

There is another method in which one listens for noise, but this time the meaning is more conventional. The technique "listens" to the sound generated by the breakdown of materials. Absence of a commercially available apparatus limits its application at this time.

Further Reading

1. K. Uosaki and H. Kita, *J. Electroanal. Chem.* **259**: 301 (1989).
2. J. W. Schultze (organizer), The Technology of Electrochemical Micro Systems, University of Dusseldorf, 1996.

12.13. A BIRD'S EYE VIEW OF CORROSION

12.13.1. Description

Civilization depends on the protection of metals, for most of them are unstable in normal environments unless they are protected by some kind of oxide film. The basic idea about the theory of corrosion is that the metal gets involved in a kind of local fuel cell in which it consumes itself. The partner to most of this self-dissolution is the deposition of hydrogen (favored in acid solutions) or the reduction of oxygen (favored in alkaline). Corrosion is measured in many ways, but the quick way in the laboratory is to move the potential a little bit away (~5 mV) from the corrosion potential in both anodic and cathodic directions and measure the corresponding current. A simple equation takes the data from this type of measurement and produces the corrosion rate.

Pourbaix diagrams relate the reversible potential of a substance suspected as a cause of corrosion to the reversible potential of hydrogen evolution or, alternatively to that of the oxygen reduction reaction, with all three plotted as a function of pH. Insofar as the metal suspected of corrosion has a reversible potential negative to those of either the hydrogen evolution or the oxygen reduction reaction (these latter potentials vary with pH), it is in a thermodynamic state favorable to corrosion. The Pourbaix diagram allows an instant determination of whether (on the above criterion) there is any possible danger of corrosion for any system examined at a given pH. This

information is useful, but it does not tell the operator whether the specimen will corrode at a *significant* rate.

A better method, from the point of view of fundamentals, is to plot the log of the current densities of the anodic dissolution current and that of the cathodic partner reaction as a function of potential, but at a given pH, respectively. The common log i at which they intersect determines the corrosion rate. These Evans–Hoar diagrams are fundamentally correct and tell whether the corrosion will be significant. However, the relevant data, which would have to take into account the presence of oxide films, etc., is at present sparse, so that Evans–Hoar diagrams are largely of value for teaching principles and seldom for giving industrially useful information on demand.

One of the most important reactions in materials chemistry is that of the dissolution of iron. However, the mechanism of this is rather complex because the rate-determining step changes with conditions, even, e.g., with stress (which may be very localized in real-life specimens). On the other hand, the dissolution mechanisms often involve an intermediate such as $FeOH^+$, which allows some guidance in the following complex traffic.

Corrosion occurs wherever there is a metal, the thermodynamic potential of which (on the normal hydrogen scale) is negative to that of some partner reaction. This chapter contains many graphic (and some unexpected) examples of practical corrosion, e.g., drops of moisture condensing on a metal surface tend to stimulate corrosion in the metal underneath the drop.

Stopping corrosion is of course the objective of the corrosion engineer. To summarize, the methods for preventing corrosion can be divided into two groups. In one, the situation is exemplified by that of a ship at sea—a metal in an infinitely large solution. To protect a metal from corrosion in such a situation, one makes the metal a cathode in an electrochemical cell (the anode is tugged along with the ship) and polarizes the ship's hull in the negative direction with respect to the corrosion potential, so that the corrosion rate (which in any case is kept very low by protective paint) is reduced. In the other group of methods, the object to be protected is in a limited amount of solution and a corrosion "inhibitor" (a chemical substance) is added to the solution. This is often an organic compound, and most of the inhibitors that have been developed empirically over the years are large, complex molecules, mostly having an aromatic component. Their most important structural characteristics and how they work have been described as well as methods for making them environmentally friendly. This latter is particularly necessary when they are used to inhibit corrosion of pipes leading to oil rig platforms located in the sea. If the inhibitors are toxic, they will damage sea life.

Inhibiting the corrosion of aluminum alloys by adding 1–5% of transition metals is a dramatic case of corrosion protection because of the small amounts of additives that are successful in reducing the corrosion rate by 1–2 orders of magnitude. It turns out that the alloying materials shift the pzc toward the positive side on the potential scale. Thus, in many practical situations, the alloys of the transition metals are in a

potential region negative to the potential at which Cl^- adsorbs and hence remain successfully uncorroded because the surface oxide film they contain is not attacked by aggressive ions such as Cl^-.

One of the methods used in corrosion control is "anodic inhibition." The method applies in particular to iron and its steels. The electrode is moved in the anodic direction (at first stimulating the corrosion rate), but soon an oxide film forms and reduces the dissolution current. There are certain types of oxide film, passive films, that are particularly protective. Indeed, such films are involved in the way metals preserve themselves in nature. There is much to be found out about these films (why they are so protective) and some of the material that allows us to understand them and their eventual breakdown by aggressive ions such as chloride, has been given in this chapter.

Corrosion in practice is usually a local matter. The idea first proposed—that one is dealing with a flat plate, all parts of which dissolve at the same time—seldom applies in practice. In the real world, corrosion nearly always occurs in tiny patches. Pitting and crevice corrosion and the invisible spreading of cracks throughout a metal are extremely important examples of corrosion. Such types of corrosion can cause dramatic, dangerous situations. The metal may look all right on the outside, but it may be developing unseen cracks inside. It was only in the latter half of the twentieth century that it was realized that the most important corrosion is invisible corrosion and that in extreme cases it can cause the collapse of bridges or even (occasionally) the breakup of ships stressed in storms.

Hydrogen is found everywhere in aqueous electrochemistry, and it plays a large part in materials science, often taking part in the mechanism of the breakdown of materials. Unfortunately, the materials that are most used in engineering construction, iron and its alloys (steels), are susceptible to the diffusion into them of hydrogen, which under certain circumstances will cause a catastrophic loss of strength of the material.

Hydrogen diffuses preferentially to sites of stress in metals, as was hypothesized many years ago by Troiano at the Case Western Reserve University in Cleveland, Ohio. This preference of hydrogen for points where the stress is greatest also throws a light upon how it takes part in stress-corrosion cracking. The latter is a troublesome and complex process that consists of two types of destructive motion. The one is understandable at once from knowledge of the electrochemistry of corrosion obtained so far. A combination of anodic dissolution at the bottom of a crack and a coupled cathodic component further up the crack or at the external surface of the specimen causes the crack to slowly advance in an entirely electrochemical way. However, in some cases the crack undergoes a sudden jump and it travels more in 1 s than in the preceding several hours. This is caused by the fact that hydrogen gathers under the stress at the bottom of the crack and when its concentration is sufficient (made higher locally by the stress there), the surrounding metal becomes weak enough for the stress to mechanically *tear* the metal and cause the crack to advance rapidly through the section of metal weakened by H damage. After this, it stops and slows into electro-

chemically controlled movement again, starting the buildup of H damage at the new crack bottom, followed after some time by another "tear," etc.

Hydrogen gets trapped inside voids in metals, and this can cause high local pressures, even thousands of times more than those in the solution in contact with the surface. Until recently, these pressures had been theoretically indicated but not experimentally demonstrated, but in this chapter two methods for measuring them were described.

A curious phenomenon in metals is fatigue. When a metal is rhythmically stressed for a sufficient time, it eventually undergoes a sudden weakening that leads to breakdown. This collapse is greatly speeded up when the metal is in contact with an electrolytic solution, such as sea water. Theories of fatigue are still rather speculative, but two are mentioned in this chapter. The effect can be dangerous in aircraft, where rhythmic stress has led to damaging situations. It is a matter for concern that aircraft exposed to certain smoggy conditions on the ground may have condensed upon them raindrops of pH 3 (Simnad, 1992), a dangerously acidic and corrosion-stimulating condition.

The rest of the chapter has been devoted to "special topics" and in materials science there are many possibilities. Those selected include the mechanism of the flotation of minerals in which the addition of a certain organic to the solution causes a specific mineral to become hydrophobic so that it is exposed to air bubbles, the bubbles stick to it and buoy the mineral up to the surface, leaving unwanted minerals on the bottom of the tank. It turns out that the mechanism of this phenomenon involves a mixed-potential concept in which the anodic oxidation of the organic "collector," often a xanthate, allows it to form a hydrophobic film upon a semiconducting sulfide or oxide, but only if there is a partner reaction of oxygen reduction. This continues until there is almost full coverage with the dixanthate, and the surface is thereby made water-repelling.

The last part of the chapter refers to the use of techniques that are relatively new. They consist of scanning tunneling microscopy and atomic force microscopy by which one can get better than 10 Å resolution in looking at surfaces. A laser beam, in collaboration with a quadrupole mass spectrometer, can be used to detect local hydrogen with a resolution of about 0.1–1 μm. Finally, ideas of electronic noise are applied to the random fluctuations of the potential of an electrode and can give a value of the corrosion current without any influence of the IR drop. This has made the method particularly useful for nonaqueous solutions.

The damage caused by corrosion costs several percent of the gross national product of technologically advanced countries! Nowhere else in the whole of science could so much money be saved by the application of the work of so few.

Further Reading

1. H. J. Flitt, J. Pezy, and J. O'M. Bockris, *Int. J. Hydrogen Energy* **8**: 39 (1983). The neodynium Yag laser and the detection of H in local areas.

2. K. C. Pillai and J. O'M. Bockris, *J. Electrochem. Soc,* **131**: 568 (1984). The mixed-potential theory of separative mineral flotation; a quantitative study.
3. K. C. Pillai and V.Y. Young, *J. Colloid Interface Sci.* **103**:103 (1985). X-ray photoelectron spectroscopy study of xanthate adsorption on pyrite mineral surfaces.
4. F. Fen and A. J. Bard, *J. Electrochem. Soc.* **136**: 166 (1989). Scanning tunneling microscopy and the corrosion of stainless steel.
5. R. Sonnenfeld, J. Schneir, and P. Hansma, in *Modern Aspects of Electrochemistry,* B. E. Conway, R. H. White, and J. O'H. Bockris, eds., Vol. 21, p. 1, Plenum, New York, 1990. STM in electrochemistry.
6. R. C. Bhardwaj, A. Gonzalez-Martin, and J. O'M. Bockris, *J. Electrochem. Soc.* **138**:1901 (1991). Scanning tunneling microscopy and the corrosion of iron.
7. J. P. Thomas and R. P. Wei, *Mat. Sci. Eng.* **A159**: 205,233 (1992). Fatigue in metals.
8. R. Woods, in *Modern Aspects of Electrochemistry,* R. H. White, J. O'M. Bockris, and B. E. Conway, eds., Vol. 29, p. 401, Plenum, New York (1996). The mechanism of the separation of minerals by means of preferential flotation.
9. G. C. Farrington, K. Kowal, J. De Luccia, J. Y. Josefowicz, and C. Laird, *J. Electrochem. Soc.* **143**: 2471 (1996). Atomic force microscopy in the corrosion of alloys.
10. A. Michaelis and J. W. Schultze, *Thin Solid Films* **274**: 82 (1996). Microellipsometry in corrosion.
11. F. Mansfeld, C. C. Lee, and G. Zhang, *Electrochim. Acta* **43**: 435 (1998). Comparison of noise and impedance data.

EXERCISES

1. In acid solutions, hydrogen evolution and in alkaline solutions oxygen reduction are, respectively, the usual cathodic partner reactions in the corrosion of metals in a solution in contact with air. Explain this in thermodynamic and kinetic terms. (Bockris)

2. The two cathodic partner reactions in corrosion are hydrogen evolution and oxygen reduction. Consider the Electrochemical Series (for a full list, see *The Handbook of Chemistry and Physics*) and work out a rule that gives the standard reversible electrode potential, less negative than which (pH 7 and $a_M zt = 10^{-6}$ *M*) a metal will no longer have a tendency to corrode (a) in 1 *M* acid and (b) in 1 *M* alkali. (Bockris)

3. Answer the following questions concerning anodic protective films:

   (a) What is the passivation potential? (Draw illustrative diagram). (b) What is the Flade potential? (Draw illustrative diagram), (c) Briefly describe the characteristics of a passive layer on a ferrous metal. (d) Give the evidence that nonstoichiometry is an essential aspect of the structure of such films, (e) What part does water, or species from water, play in forming the structure of a passive

film? Review the evidence for the importance of water both for Fe and Al films. (Bockris)

4. (a) How do corrosion pits begin? (b) By what experimental means can the interior of individual pits be studied? (c) The interiors of pits in ferrous metals tend to have a pH around 4, independent of the external pH. (d) What molecular happenings would lead to such a result? (A quantitative answer is expected.) (Bockris)

5. (a) Draw a diagram illustrating crevice corrosion. (b) Where do the cathodic reactions occur and which of the two usual cathodic reactions is considered here? (c) What is the essential reason that the danger of corrosion in crevices is greater than that on the surface? (Bockris)

6. Pipelines (e.g., carrying oil) are subject to corrosion. They pass through various kinds of ground, some in claylike soil and others in more sandy soil. It is found that a pipe in a claylike soil corrodes more readily than that in contact with a sandy soil. Explain this. Draw a diagram on which you point out the influence of O_2 and pH on the corrosion of the steel pipe under the conditions described. (Bockris)

7. According to literature data, the corrosion rate of Fe in 3.5% NaCl saturated with O_2 under a pressure of 1 atm is 9.3 mm yr^{-1}. Estimate the magnitude of the corrosion current density of Fe. (Gokjovic)

8. Estimate the magnitude of the corrosion current density and corrosion potential of zinc in 0.01 M H_2SO_4 using the following data:

TABLE E.1

Electrochemical Reaction	Solution	i (mA cm^{-3})	α
$Zn^{2+} + 2e^- = Zn$	1 M ZnSO$_4$	11	1.5
$2H^+ + 2e^- = H_2$	0.5 M H$_2$SO$_4$	3.2×10^{-6}	0.5

The H_2 evolution reaction on Zn is first order with respect to H^+ ions. Assume that the concentration of Zn^{2+} ions in solution of H_2SO_4 in which Zn corrodes is 10^{-6} M. Estimate the magnitude of the corrosion current. (Gokjovic)

9. What current has to be supplied by an external source to protect 1.00 m^2 of Zn from corrosion in 0.01 M H_2SO_4? Take your data from exercise 8. (Gokjovic)

10. The cathodic reaction during corrosion of iron in sea water is oxygen reduction. Solubility of O_2 from the air in sea water is 0.189 mol m^{-3} and the diffusion coefficient of O_2 is 2.75×10^{-9} m^2 s^{-1}. The diffusion layer thickness in an unstirred solution is about 0.5 mm. (a) Estimate the corrosion current density of iron in sea water. (b) If iron is connected to the negative pole of an external

power source, what current density has to be applied to corroding iron to slow it down by ten times its corrosion rate? (Gokjovic)

11. Estimate the corrosion potential E_{corr} and the corrosion current density i_{corr} of Zn in a deaerated HCl solution of pH 1 at 298 K. In this solution Zn corrosion is accompanied by the hydrogen evolution reaction (h.e.r.). The parameters (standard electrode potential $E°$, exchange current density i_0, Tafel slope b of Zn dissolution and the h.e.r. on Zn are

$$E_{Zn}^{\circ} = -0.76 \text{ (vs. SHE)}, \; i_0 = 3 \times 10^{-1} \text{ A m}^{-2}, \; b_{anodic} = 0.030 \text{ V/dec}$$

$$\text{At } E - E_{H+/H_2}^{eq}, \; i_0 = 1 \times 10^{-7} \text{ A m}^{-2}, \; b_{cathodic} = 0.120 \text{ V/dec (Numata)}$$

12. A polarization study of Fe corrosion in an alkaline sulfate melt containing $Fe_2(SO_4)_3$ under SO_2-O_2 at 973 K reveals that:

Anodic reaction	$Fe \rightarrow Fe^{2+} + 2e$
Cathodic reaction	$Fe^{3+} + e \rightarrow Fe^{2+}$
Chemical reaction at metal or oxidation interface.	$2Fe^{2+} + SO_3 + 1/2O_2 \rightarrow SO_4 2 + 2Fe^{3+}$
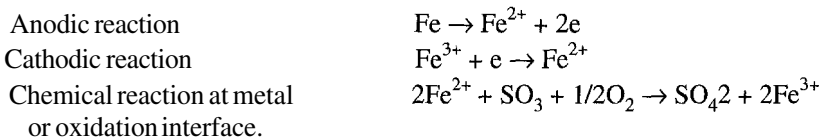
The corrosion current density i_{corr} is evaluated by the electrochemical polarization resistance method assuming that both the anodic and the cathodic partial currents obey the Tafel relation:

$$i_{corr} = RT/R_p F(\alpha_{anodic} + \alpha_{cathodic})$$

where R_p is the polarization resistance obtained as the reciprocal (dE/di) of the slope of the voltammogram (polarization curve) in the vicinity of E_{corr} and α_{anodic} and $\alpha_{cathodic}$ are the transfer coefficients of the partial anodic and cathodic currents. The values of α_{anodic} and $\alpha_{cathodic}$ are given by the Tafel slopes of the anodic and cathodic currents b_{anodic} and $b_{cathodic}$.

$$\alpha_{anodic} = 2.30 \, RT/b_{anodic}F, \; \alpha_{cathodic} = 2.30RT/b_{cathodic}F$$

(a) Calculate the corrosion current density, given that b_{anodic} and $b_{cathodic}$ are 0.340 and 0.180 V/dec, respectively, and R_p is 2.05×10^{-4} Ωm^2.

Measurement of mass lost is the conventional method for determining the corrosion rate. The mass loss of an Fe specimen immersed in a corrosion test potential is determined by weighing. (b) Convert the mass loss rate 2.34×10^2 g $h^{-1}m^{-2}$ into i_{corr} using the atomic weight 55.847. (c) What is the difference between the results of the mass loss measurement and the polarization resistance measurement? (Numata)
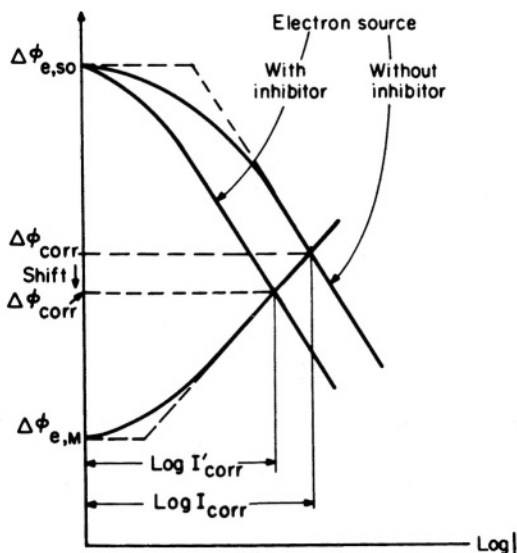
Fig. E12.1

13. The thermodynamic stability of tin is studied when the metal is introduced in an aqueous solution of pH 7 at 25 °C under two conditions: (a) the amount of dissolved oxygen is that corresponding to an atmospheric pressure of 0.20 atm, and (b) the solution is deoxygenated by the passage of nitrogen gas. Write the reaction taking place in each instance, draw the cell diagram, and calculate the potential difference for the free oxygen solution, indicating the physical state of the metal, i.e., inhibition, passivation, or dissolution. Consider:

$$E^\circ_{Sn^{+2}/Sn} = -0.136 \text{ V}, \ E^\circ_{CO_2/H_2O} = 1.229 \text{ V}, \ a_{Sn^{2+}} = 10^{-6} \text{ and } p_{H_2} = 1\text{atm}$$

(Zinola)

14. With the help of the Evans diagram (Fig. El2.1) explain the influence of an inhibitor on the corrosion potential and corrosion current. Assume that the inhibitor decreases the exchange current density for the cathodic reaction. (Contractor)

15. The anodic and cathodic reactions occurring during the electrolysis of a solution of $TlNO_3$ in HNO_3 are:

$$Tl^+ \rightarrow Tl^{+3} = 2e^- \text{ and } Tl^{+3} + 2e^- \rightarrow Tl^+$$
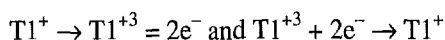
For a cell built with separated compartments, the resistance of the solution and the joint is $0.5\ \Omega$ at 25 °C. The area of both electrodes is $1.0\ m^2$ each. The electrolysis of the solution gives the following results.

TABLE E.2

$i\ (mA\ m^{-2})$	$\eta_{anodic}(mV)$	$|\eta_{cathodic}|(mV)$
0.19	32	24
0.27	61	43
0.33	93	64
0.45	124	86
0.56	156	101

Determine the maximum rate of formation of the T1 (III) salt in mol m^{-2} s if the difference of the applied potential is 1.0 V. In this case it is possible to neglect the decomposition reactions of the solvent. (Zinola)

16. A piece of iron corrodes in an aqueous solution free of oxygen and saturated with hydrogen at 1 atm. The pH of the solution is 3.1 and the activity of the ferrous ions is that corresponding to a concentration of 0.02 M. For the electro-oxidation of iron it is known that $i_0 = 9 \times 10^{-7}\ A\ cm^{-2}$, and $(\partial\eta/\partial \log i)_T = 0.04$ V dec^{-1}, and for the hydrogen evolution reaction $(\partial\eta/\partial \log i)_T = -0.12$ V dec^{-1}. (a) Determine the corrosion rate of iron if the corrosion potential is –0.398 V vs. NHE at 25 °C. (b) Calculate the exchange current density for the hydrogen evolution reaction on iron. Neglect any effect due to ionic activity and mass transfer. Consider that $E^{\circ}_{Fe^{+2}/Fe} = -0.447$ V. (Zinola)

PROBLEMS

1. Using thermodynamic data available in *The Handbook of Chemistry and Physics*, draw the potential–pH diagrams for a hydrogen and for an oxygen electrode. In each case, assume that the relevant gas is present at $p = 1$ atm. In addition, draw the potential–pH diagram for Mg. As a simplification, assume that in a solution of pH < 1, $C_{Mg^{2+}} = 10^{-6}\ M$. Write in your diagram a clear indication of the region of pH and potential where corrosion occurs, where the metal is passive, and where it is immune to corrosion (construction of a Pourbaix diagram). (Bockris)

2. It is well known that Pourbaix diagrams give the thermodynamic limits of corrosion. However, it is possible that corrosion in a system may be limited by kinetics to rates so low that corrosion that is thermodynamically possible can be neglected under practical circumstances. In this light, (a) construct an Evans diagram, i.e., a plot of the actual relevant electrode potentials against log i for

two reactions: one, the dissolution of Ni and the other, the cathodic evolution of H_2. Assume that i_0 for the dissolution of Ni (into a solution containing $a_{Ni^{2+}} = 10^{-6}$ M) is 10^{-8} A cm^{-2} and $b_{anodic} = 0.038$, for pH 0. Assume that $(i_0)_{H2} = 10^{-4}$ A cm^{-2} and $b_{cathodic} = 0.116$. Other relevant data are available in tables.

(b) Determine the corrosion potential and the corrosion rate (in amperes cm^{-2} and then in microinches per year). (Bockris)

3. Consider an Evans diagram in a general way. The anodic dissolution reaction is to be represented in the Tafel region; the same applies for the cathodic partner reaction. (a) Draw the two Tafel lines and show the region of intersection $(i_{anodic} = i_{cathodic})$. Indicate on the graph the corrosion rate and corrosion potential.

(b) Write the expression for the net anodic current as one biases the electrode away from the corrosion potential. The overpotential term in the Butler–Volmer equation can be replaced by V-V_{corr}.

(c) Now consider the anodic current thus obtained [the i_0 will be $(i_0)_{corr}$] and linearize it for low overpotential. Derive from this the Stern–Geary expression.

$$i_{corr} = \frac{RT}{F(\alpha_{cathodic} + \alpha_{anodic})} \times \frac{di}{dV}$$

where α_{anodic} and $\alpha_{cathodic}$ are Tafel parameters (transfer coefficients). (Bockris)

4. Cathodic protection is an effective method of reducing corrosion when the object to be protected is in a solution in which the use of corrosion inhibitors is impractical (e.g., sea water). However, in cathodic protection systems working under potentiostatic control, it is important to restrict the cathodic potential to avoid an evolution of more H_2 than is necessary to reduce the corrosion rate to negligible proportions (because H contacting the steel hull of a ship through fissures in the paint may introduce embrittled areas in the hull).

(a) Draw an Evans diagram showing the Tafel line for Fe dissolution at pH 7 in a Cl$^-$ solution. (Tafel slope $RT/2F$) and $(i_0)_{Fe} \simeq 10^{-8}$ A cm^{-2}. (b) Draw a cathodic hydrogen evolution (10^{-6} for i_0 and $b_{cathodic} = 0.116$). (c) Show the intersection of the two lines at i_{corr} and then extrapolate the hydrogen line further until a potential has been reached at which the Fe dissolution rate has been reduced 1000 times compared with the spontaneous corrosion rate. What current density of hydrogen evolution is reached? (d) By examining the conditions for embrittlement in Section ( ), find whether this cathodic "protection," in fact, endangers the object to be protected because it embrittles the underlying metal. (Bockris)

5. Conditions inside pits are complicated by the aggregation of hydrous oxides and the effect of tensile stress, which increases the dissolution velocity (per unit area) at the tip over that of a plain sheet of metal. A zeroth approximation for the electrochemistry of corrosion in a pit is to take the hemispherical tip as anodic

and dissolving. The cathodic partner for this is often oxygen reduction on the surface of the metal. (In reality, the dissolving area spreads up the pit from the tip and the area in which reduction occurs enters the pit from the exterior.)

(a) Draw an Evans diagram for the rate of advance of the pit. Make the following zeroth approximation: the cathodic reduction of O_2 from air occurs on 1 cm^2 of external surface only. The anodic dissolution takes place on the surface of an inverted hemisphere (the bottom of the pit), the radius of which is 10 nm. In the Evans picture one must take into account the gross difference in the size of the anodic and cathodic areas and the stress effect on i_0 for Fe dissolution (take as 10^{-5} A cm^{-2}); that for O_2 dissolution on passive Fe is 1.3×10^{-7} A cm^{-2} at pH 7. The reversible potentials are available in handbooks.

(b) Express the rate of growth of the pit (neglecting any non-Galvanic "cracking") in millimeters per hour. (Bockris)

6. Local stress is a potent cause of corrosion and may occasionally lead to distressing events, e.g., the falling off of a car door. Metals extracted from their ores may have been at one stage in a liquid state. They cool and solidify, thus introducing local stress points. Further, even in a solid that has not been subjected to stresses in its treatment, stress develops at its dislocations.

Consider a bar of iron subject to an external (tensile) stress of 10 kg mm^{-2}. In this bar there are some corrosion pits. The surrounding atmosphere is moist and acidic. Consider a pit 10^4nm deep ($= l$) with a tip having a radius, r, of 10 nm. A rough formula for the initial stress at the tip of a pit is that the external tensile stress there is increased $2l/r$ times compared with the tensile stress applied to the specimen.

In Fe, the solubility of H is given by about 7×10^{-9} mol cm^{-3} at 1 atm of H_2 pressure. Calculate its value at the tip of the pit mentioned, assuming the partial molar volume of Fe is 2.4 cm^3 mol^{-1}. Comment on any deductions you might wish to make about the movement of the tip in Fe containing a much enhanced degree of H.

7. Stress corrosion cracking is an insidious phenomenon, invisible to the naked eye, and is responsible for disastrous happenings: the collapse of the snow-laden roof of a stadium the falling down of a steel bridge, etc. As with most practical corrosion, it is caused by several factors. Discuss the part played by the following factors: (a) enhanced dissolution velocity at the stressed crack tip, (b) permeation of H to the crack tip, (c) cracking of a protective passive layer, and (d) "tearing" of the metal weakened by H absorption.

In assessing the relative role of these factors under varying ambient conditions (e.g., the presence or absence of aggressive anions such as Cl$^-$), recall that all metals have a degree of strain (e.g., fraction elongated) at which they will lose the ability to resume their original structure when the stress is released. Nonelas-

tic deformation leads to breakdown, independently of any of the mechanisms of corrosion.

8. One of the long-recognized causes of loss strength by metals is H embrittlement. A "high pressure in voids" mechanism to explain this was first suggested by Zapfe and Sims in 1941. Their concept (that gigantic pressures build up inside voids in a metal due to diffusion of H from the surface) seems at first thermo-dynamically unacceptable because there is usually an outside pressure of H_2 in an aqueous solution in contact with the metal that is equivalent to 1 atm of H_2; therefore at equilibrium, it is not possible that in the voids the H_2 pressure (fugacity?) could exceed 10^4 atm. The situation may be rationalized, however, by accounting for hydrogen overpotential (i.e., irreversibility) in the reactions of H at the interface of the metal in contact with the aqueous phase. This phenomenon destroys the idea of thermodynamically reversible equilibrium between the H on the electrode surface and any H_2 that may collect inside voids within the metal (thus potentially causing the spread of voids and tending to embrittle).

    It turns out that the pressure developed in the voids depends upon the mechanism by which H desorbs from the metal/solution interface to form H_2 (which depends not only on the metal but also on difficult-to-control factors such as the adsorption on the surface of surrounding organic debris in real solutions).

    (a) Calculate the overpotential at which, with $H^+ + H_{ads} + e \rightarrow H_2$ rate determining, the metal may be in danger. (b) If Fe is in contact with an 0.1 M HCl, construct an Evans diagram to assess the corrosion potential in the H_2 scale. The parameters for hydrogen evolution and iron dissolution are given above. (c) Calculate the hydrogen overpotential during hydrogen evolution corrosion and compare the resulting p_{H_2} in voids in the interior of the metal with its spreading pressure. (d) Decide whether corrosion here will also cause H embrittlement. (Bockris)

9. Metals are obtained by the treatment of oxides and sulfide ores found in the earth. However, there is an initial difficulty—the desirable ores are often mixed up with those of little commercial value, and the problem is to obtain the desired ore free from those of lesser worth. For many years now, largely due to the initiative of Australian workers, it has been possible to find organic substances which, when added to a suspension of mixed ores, pick out the desired one, and (when air is bubbled into the system) float it to the surface, from which it can be "raked" off, i.e., separated and made available for chemical or electrochemical extraction of the metal. It turns out that the basis of this mineral flotation technology involves the Wagner and Traud mixed-potential concept and is thus indirectly related to corrosion theory.

    (a) What are typical organic substances that perform this economically successful process? (b) What do they actually do on the surface that is relevant

to ideas on the mechanism of flotation? (c) What bubbles are adsorbed and what condition is necessary before they can "stick" on the surface of the desired ore particles and float then upward? (d) How is mixed potential involved and what are the two reactions that form the mixed potential? (e) In what way are the electrocapillary phenomena (Section) involved? (Bockris)

10. Sea water has the following composition (mg/liter)

TABLE P.1

Chloride	19,860
Sodium	10,460
Calcium	420
Magnesium	1310
Bicarbonate	150
Sulfate	3230

The pH is about 8.2 and the amount of oxygen contained in sea water is about 6 ppm. Most metals are corroded in contact with sea water. (a) What type of corrosion occurs in sea water? (b) What technique can we use to inhibit the corrosion due to sea water? (Pou)

11. In the refinery, crude oil is heated at 350 °C and passed through an atmospheric distillation column. The distribution of temperature of the column should be about 350 °C in the bottom and 50 °C at the top. The crude oil contains dissolved gas, H_2S, for example, and a small amount of water that contains dissolved $CaCl_2$ and $MgCl_2$. (a) What type of corrosion problems should we expect? (b) What type of protection should be adopted? (c) What type of corrosion inhibitor should be used? (Pou)

12. Anaerobic bacteria living in water can damage a metal in an oilfield or gasfield. These bacteria can grow under the conditions encountered in a reservoir or a pipe. The most famous are the sulfate-reducing bacteria, known as SRB. They use sulfate as a nutrient and transform it into sulfur according to:

$$SO_4^{2-} + 8H \rightarrow 4H_2O + S^{2-}$$
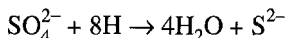
H is obtained from the enzyme hydrogenase.
   (a) In an oil and gas field, if the water is acidic, what happens to the metal? (b) Write the corrosion reaction. (c) What is the best way to protect metal from the corrosion caused by SRB?

13. In a natural gas field, we produce natural gas, condensate, and water. The gas composition is (in mol%):

TABLE P.2

Nitrogen	23.32%
Hydrogen sulfide	10.95%
Carbon dioxide	nil
Methane	63.2%
Ethane	0.35%
Propane	0.06%
isobutane	0.04%
n-Butane	2.03%
isopentane	0.03%
n-Pentane	0.01%
Hexane and heavier	0.01%
Water: condensation water	1 g/liter of NaCl

The surface equipment, pipe, and tubing are of carbon steel. The wall shear stress of the liquid due to the high gas velocity is $90 \ N/m^2$. (a) What type of corrosion is encountered? (b) What type of protection can be adopted economically? (c) If a corrosion inhibitor is adopted, what type of selection tests should be used to choose the right chemical? (Pou)

14. A piece of iron of a real area of $1 \ cm^2$ is placed in a solution containing 0.5 g-ions dm^{-3} of Fe^{2+} (mean activity coefficient = 0.46) at pH 0, adjusted from H_2SO_4 at 298 K and $p_{H_2} = 1 atm$. (a) Without paying attention to any refined mechanism and using only those data plus general constants; use the reasoning of a Pourbaix diagram to find out of the ion is stable or corroding. (b) Estimate the corrosion affinity (cell potential), the free energy change, and the corrosion potential. (c) Does the process evolve spontaneously? (d) Calculate the corrosion flux ϕ_{corr} (rate in terms of equiv. $g^{-2} \ s^{-1}$, the corrosion rate v_{corr} ($g \ cm^{-2} \ s^{-1}$), and the penetration rate p_{corr} ($mm \ yr^{-1}$) if the corrosion current density i_{corr} is $10^{-3.83}$ ($A \ cm^{-2}$), $A_{Fe} = 56$, and the metal iron density p is $8 \ g \ cm^{-3}$. (Plonski)

15. The 75 values listed under E_{corr} and pH in Table P.3 are steady-state corrosion potentials as a function of pH, collected from the literature.

(a) Select the E_{corr} values measured at pH 0.3, establish the midpoint of E_{corr} values, and plot and analyze the histogram. (b) Are you faced with random errors that can be treated by statistical methods, or are your values affected by some systematic errors (bias)? (c) After excluding the values that are too far from the midpoint, calculate the mean value, E_{corr}, and its estimate of standard deviation. (d) If a good reproducibility means an estimate of standard deviation $S_Y < 6 \ mV$, what do you think about your results?

(e) Plot all the $n = 75$ pairs of E_{corr} and pH values given in Table P.3. (f) Observe whether the plot shows a range of acceptable corrosion potential values at each pH. (g) Fit the data to a linear plot of the form $Y = aX + b$, using the method of least squares. In this case $Y = \hat{E}_{corr}, b = \hat{E}_{corr}(pH \ 0;$

$X = \text{pH}$; $a = \Delta\hat{E}_{corr}/\Delta\text{pH}$, where the symbol "" refers to the theoretical least-square straight line. (h) Plot the theoretical least square line.

(i) Using the resultant least-square line established in (j), calculate the standard deviation of this line, S. Plot on the same graph the $2S$ limit lines that encompass the least-square line, i.e.,

$$\hat{E}_{corr} = (\hat{E}_{corr} (\text{pH } 0) \pm S_E) + (\Delta\hat{E}_{corr}/\Delta\text{pH})\text{pH}$$

TABLE P.3
Steady-State Corrosion Potential, E_{corr} (mV/NHE)
vs. pH for Iron in SO_4^{2-} and ClO_4^{-} Solutions
Reported by Different Authors (Plonski)

pH	$-E_{corr}$	pH	$-E_{corr}$	pH	$-E_{corr}$
0.0	232	0.4	295	2.0	326
	240	0.5	250		330
	250		259		340
	260		274		364
0.1	273	0.55	270		382
	280	0.8	260	2.5	330
0.30	224		293		330
	225		320	2.7	348
	235	1.0	293		358
	245	1.2	278		390
	248		282	2.8	392
	250		305		358
	251	1.3	270		420
	253	1.4	290		422
	255	1.5	297	2.9	362
	260		318		370
	260		333		422
	260		384	3.1	397
	268		386	3.3	397
	268	1.9	325	3.5	430
	273		327	3.7	400
	273		342		415
	276		363		434
	277		383	4.0	435
	279				450
	280				

16. It is considered that the slope $\Delta E_{corr}/\Delta\text{pH}$ has mechanistic significance. Various corrosion mechanisms for iron in acid media have been proposed in the literature. One of them (the catalytic mechanism) is based on the value $\Delta E_{corr}/\Delta\text{pH} = -47$ mV; another (stepwise electron transfer mechanism) is based on the value -59 mV. (a) Using the values in Table P.3 calculate the estimate of standard deviation of the slope, S_{S1}, and establish if it is low enough to discriminate between the two mechanisms above.

It is considered that the reaction order in OH^{-}, $V_{OH} = \Delta \log i_{Fe\,cathodic}/\Delta\log$

a_{Fe}^{2+} and the cathodic Tafel slope, $b_{cathodic} = \Delta E/\Delta \log i_{Fecathodic}$, have mechanistic significance. The theory predicts for the catalytic mechanism $V_{OH} = 2$ and $b_{cathodic} = -30$ mV, and for the stepwise electron transfer mechanism $V_{OH} = 1$ and $b_{cathodic} = -120$ mV. (b) Using the experimental data for iron cathodic deposition from acidic solutions given in Table P.4, establish whether these data can be used as mechanistic criteria. If not, why? (Plonski)

TABLE P.4

$V_{OHcathodic}$	N	$-b_{cathodic}$ (mV dec^{-1})	N
0.05	1	30–39	5
0.30	1	40–49	1
1.0	2	50–59	0
1.4–1.6	2	60–69	3
2	2	70–79	0
		80–89	1
		90–99	0
		100–109	0
		110–119	2

[a]Number of papers, reporting a steady-state reaction order in OH^-, V_{OHcat}, and cathodic Tafel slope, $b_{cathodic} = \Delta E/\Delta \log i_{Fecathodic}$ (mV dec^{-1}) in a certain interval of values for cathodic deposition of iron in acid media free of oxygen and complexing or adsorbable anions (literature values).

17. The oxidation rate of a stainless steel is studied in an aqueous solution of pH 3.1 saturated with H_2 and free of O_2. The following data were obtained from an Evans diagram at 25 °C.

TABLE P.5

System	E°(V vs. NHE)	i_0 (A cm^{-2})	$(\partial\eta/\partial \log i)_T$ (mV dec^{-1})
Fe/Fe^{2+}	−0.437	8.95×10^{-7}	33.7
H$_2$/H$^+$	0	2.96×10^{-6}	−118.2

(a) Determine the rate of corrosion in terms of the current density for a solution of 20 mM in Fe^{-2}. Assume mass transfer effects and ohmic drop to be negligible. At one point during the experiment the pH increases and oxygen flows into the electrolyte owing to a lack of a good deoxygenation system, (b) Determine the corrosion reactions occurring in the system as well as the new rate of corrosion. Assume that the new pH of the solution is 4.5 and that the amount of dissolved oxygen creates an equilibrium pressure of 0.75 atm. Consider that for the reduction of oxygen:

TABLE P.6

System	E° (V vs. NHE)	i_0 (A cm^{-2})	$(\partial\eta/\partial \log i)_T$ (mV dec^{-1})
O$_2$/H$_2$O	1.23	5.82×10^{-9}	-141.2

(c) Write the corresponding electrochemical reactions and evaluate the new corrosion rate. (Zinola)

MICRO RESEARCH PROBLEM

1. Organic substances that act as corrosion inhibitors depend for their efficacy upon a number of compromises. Thus (see Fig.) the more negative the standard free energy of adsorption, the less negative the standard free energy of dissolution of the inhibitor. However, if the tendency to dissolve is sufficiently small, the adsorption isotherm (in spite of a highly negative adsorption standard free energy) will not cover a surface sufficiently to be an effective inhibitor.

   Consider the compromise necessary at constant potential and pH. (a) Derive an expression (assuming for simplicity a Langmuir isotherm) for the optimal relation of ΔG°_{ads} to ΔG°_{diss}.

   ΔG°_{ad} involves several factors, an important one being the tendency of atoms in the inhibitor to bond with iron (the metal for which inhibitors are most often designed). (b) By quantitative consideration in terms of the HOMO and LUMO properties of the bonds, comment on the efficiency of inhibitors containing, respectively, S and NH$_2$. (c) Correspondingly, give a molecular interpretation of the facts that most inhibitors have a significant aromatic character, (d) Deduce some rules concerning likely inhibitor structures for the protection of aluminum in acid metals. (Bockris)

CHAPTER 13

CONVERSION AND STORAGE OF ELECTROCHEMICAL ENERGY

13.1. INTRODUCTION

In the old days, say before about 1970, the title for this chapter would most certainly have been "Batteries and Fuel Cells." Note the order in which the words used to be written. In those days, batteries were well known for their use in auxiliary lighting, and above all for starting combustion engines. Fuel cells were a curiosity, worked on for more than a century by the occasional enthusiast, but unrecognized by the vast majority for their revolutionary possibilities.

All this changed in the late 1990s. It is now fuel cells and batteries, but that significant change shows only the tip of the proverbial iceberg. What caused the change? There were two important decisions, each made by an organization with great authority and weight. The first was the decision made by the National Aeronautics and Space Agency (NASA) in the late 1950s, to use fuel cells for auxiliary power in space vehicles, where so much depends on low weight per unit of energy produced (fuel cells are about three times as effective weightwise as any other method of providing electricity on board).

The second, having a far wider range of implications, was the decision in 1997 by the Daimler-Benz company of Germany (makers of Mercedes cars) to run their electric cars, not by means of electricity stored in batteries, but by means of fuel cells running on hydrogen produced by the on-board re-forming of methanol, or gasoline. The company will mass produce their electric cars for sale in 2002. Many other car makers have announced they will also use fuel cells instead of batteries for their electric cars, for this would give them a range greater than that of present gasoline-fueled cars and needing no more time for refueling than conventionally powered cars.

The words, fuel cells and batteries, lead to the misconception that there must be a close connection between the two devices. However, the connection is only apparent and the purpose of each device is utterly different. The fuel cell is an electrical energy producer (Sec. 7.1.3.2); one takes a fuel (e.g., methanol) and leads it into the oxidizing anode of a fuel cell. At the counter-cathode, oxygen in air is reduced. The free energy of the oxidation of, for example, methanol comes out not as heat (as it would in a chemical engine), but directly as electrical energy. So fuel cells *produce* electricity. They should be called *electrochemical electricity producers* (eces), although that's perhaps too much of a mouthful.

Batteries, on the other hand, must have their electricity produced elsewhere (e.g., by the present complex system of *burning* a fuel and using the heat produced to expand gases and push turbine blades, finally driving dynamos. The battery *receives* this electricity, which drives a reaction on each of two electrodes, respectively, *up* a free energy gradient for the overall reaction made from the electrochemical reaction at each of two electrodes. The "charged" battery can then be put aside and the electricity stored until it is needed. Because the battery was charged by using an outside electricity source to drive the electrode reactions uphill (against its ΔG, i.e., for a *positive ΔG*), the battery is ready to release this energy downhill in a spontaneous reaction, producing back the electricity that charged it.

So batteries store electricity and fuel cells produce it. As will be seen, their purpose and place in the economy are entirely different. As new kinds of batteries reach the commercial market, they will have important evolutionary effects on our lives, for example, they will increase greatly the length of time we can use portable computers without battery replacement, or provide long-term power for internal artificial organs.

As for fuel cells, however, *their* effects will be revolutionary because they contain within them a vital secret: They are Carnot unlimited in the efficiency of their conversion of chemical energy to electricity. In practice, this gives them the advantage of more than doubling the time over which we can use chemical fuels before our supply of them is exhausted. But fuel cells have another revolutionary advantage: They produce electricity with no polluting effluents whatsoever. With batteries one has to take into account the increased pollution from the combustion of coal or oil used to make the electricity with which to charge them. Fuel cells *and* batteries are shown in Fig. 13.1.

13.2. A BRIEF HISTORY OF FUEL CELLS

The formal discovery of the first fuel cell principle—making electricity *directly* from chemicals—is attributed to a British judge, Sir William Grove. Between his court appearances, Sir William pondered two remarkable facts. He knew that having a source of electricity—one could force water to decompose and make hydrogen and oxygen gases. However, it was known even in 1839 that hydrogen and oxygen combine with explosive zeal to make water. Water was first electrolyzed by Nicholson and Carlisle

Fuel Cells and Batteries
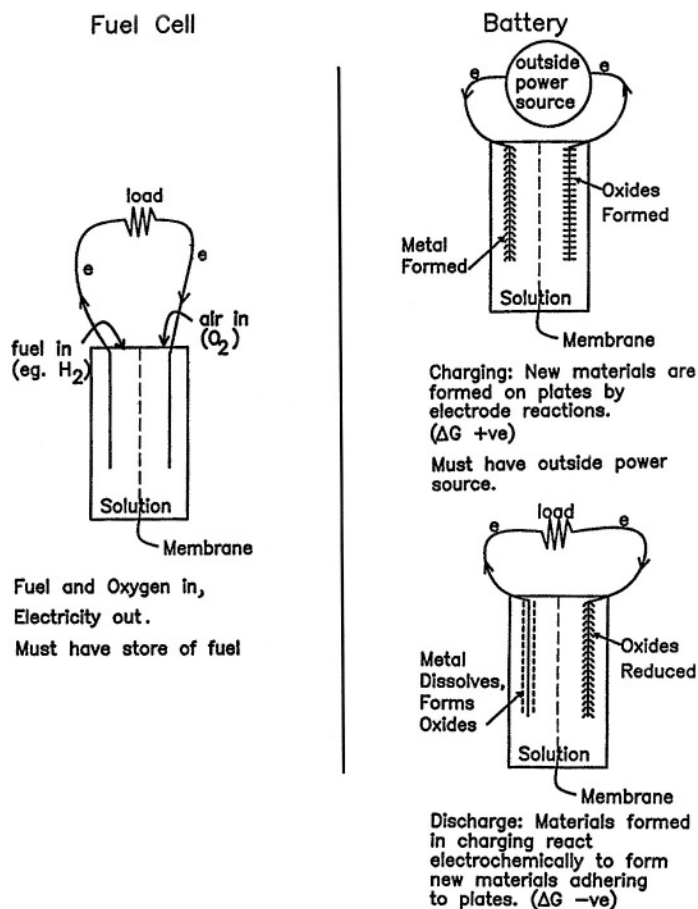
Fig. 13.1. Essentials of fuel cells and batteries: diagrammatical.

in 1800. Then a question struck Sir William: If putting electricity in made water produce hydrogen and oxygen, could putting hydrogen and oxygen into the cell make electricity?

Sir William decided that the best way to answer this would be to try it out, so he took an electrolysis cell and disconnected the electrical power source, bubbled hydrogen around one electrode, and connected the two electrodes together (the other electrode would have oxygen from the air). Sure enough, a current flowed through the wire. It kept on flowing as long as hydrogen continued to be bubbled over the electrode. Sir William was not satisfied with the puny 0.6 V that a meter showed when placed

between the two electrodes, so he put several of the electricity-producing, two-electrode cells together in series; with 50 of them each supplied with hydrogen, he produced some 25–30 V, a respectable potential (see Fig. 13.2) and a quite practical source of power.

Published in 1839, this discovery failed to impress the British, but it was recognized in Germany. Indeed, during the 1800s in Germany there were replications and improvements on Grove's seminal work, and in 1894, the renowned German chemist, Friedrich Wilhelm Ostwald, then president of the German Bunsen Society, gave a far-reaching address about energy. Reading it 100 years later, it sounds quite modern; Ostwald was a century ahead of his time. His address predicted pollution for the cities if the current path of combustion to obtain energy continued to be followed. What is more, he pointed to the great energy saving that would be attained, if instead of using the heat from the combustion of coal (the fossil fuel of the time) to expand gases, push pistons, etc., one used the direct conversion of chemical fuels to electricity offered by the fuel cell route (see Fig. 13.3).

Ostwald's ability to see the future proved to be greater than his influence on the course of events. For a while, in the early years of this century, electric cars competed with those using internal combustion, but the long charging time for the batteries proved the undoing of the use of clean electric vehicles so that the path to energy conversion in the twentieth century turned against that advocated by Ostwald and led to the pollution and inefficiency he predicted.

Fuel cells did not boom again for more than 50 years, although there was occasional activity in Europe. Then another person appeared on the scene, a man in the same vein as Sir William Grove, but much, much more persistent. This was Francis Thomas Bacon, and since it was he who stood directly behind NASA's use of fuel cells in the space flights, it can truly be said that more than any other individual, it was
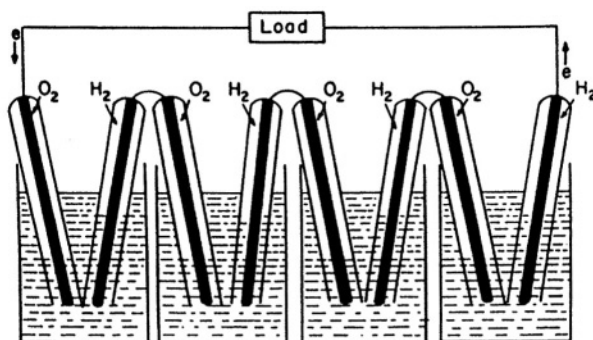
Fig. 13.2. Diagram of the first hydrogen-oxygen fuel cell, constructed by Grove (1842). The dark lines in the axes of the tubes are platinized-foil electrodes.
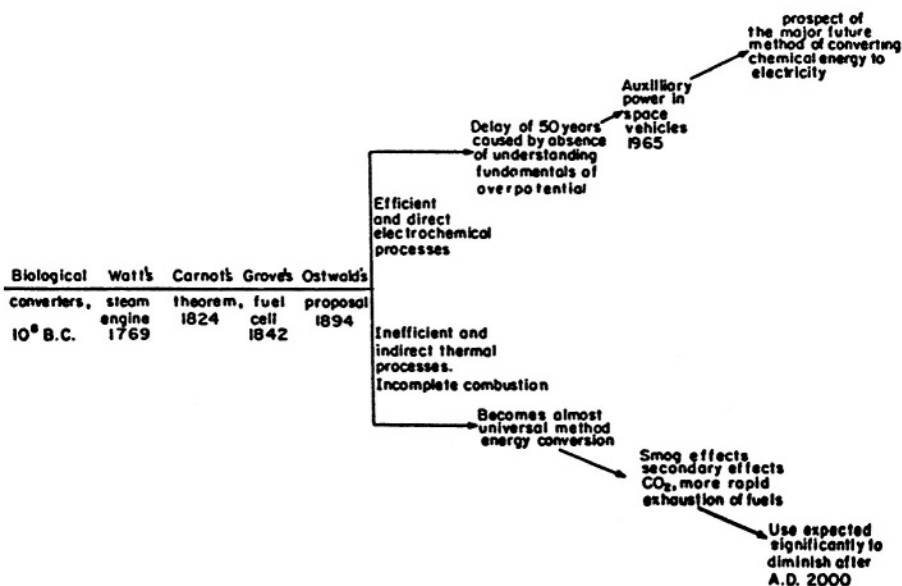
Fig. 13.3. Energy conversion took the wrong path in 1894.

he who triggered the present turn toward fuel cells as the center of energy conversion in the twenty-first century.

In 1932, Bacon was an engineer employed by a United Kingdom firm that made turbines. Totally unknowledgeable about electrochemistry, one day he glimpsed a water electrolyzer in a distant part of a company laboratory. In one of the moments often described by inventors, he rediscovered in a flash Grove's thought: electricity in, hydrogen out; so why not hydrogen in and electricity out? The turbine company would not hear of it, but Bacon decided to pursue what he thought was his discovery on his own. He made an electrolyzer, pushed in hydrogen, and indeed found electricity, as Groves had done 93 years earlier. To pursue the matter, he at first hid the forbidden project in a cupboard at work, but soon found it absorbed his whole interest. Then entered a factor without which the development of the electrochemical path to direct energy conversion might well have been delayed a further 50 years. Bacon came from a powerful family with roots back into the fifteenth century and was wealthy by inheritance. He left the turbine maker and devoted his life, and his considerable ingenuity and persistence, to making the fuel cell a practical energy-producing device.

Once more, as with Ostwald and his predictions, no one took any notice. Bacon lived in the university town of Cambridge, England and it seemed logical that the university there, famed as the site of many scientific advances, would be likely to take up his work. However, it showed not even a glimmer of interest. Undaunted, Bacon decided that if he was not to be taken into the university and given facilities there, he

would build his own laboratory in some huts on an unused airfield near the town of Cambridge and fund the work himself.

There, in the 1950s, Bacon pursued Groves' discovery. Realizing that electrochemistry must be the basis of it, in 1952 he hired Reginald G. H. Watson, a young physical electrochemist from the Imperial College of Science and Technology (Britain's MIT). With Watson doing the electrochemistry, and an engineer to help scale up the work, in August 1959 Bacon and his team finally came out with a 5-kW fuel cell which they proved could power a 2-ton capacity forklift truck.

This finally did wake up the British, so much so in fact, that a photograph was published in the London *Times* of August 1959 showing Bacon and the 5-kW fuel cell in action. A quarter of a century of work had paid off, and much better than could have been predicted from the forklift truck. This was near the time of Sputnik (1958) and the newly formed American space agency (NASA) was searching for new technology that would help it to leapfrog the Russian space initiative. Bacon's cell was at the base of the great NASA development of fuel cells, for it allowed NASA space vehicles to carry twice the stored energy per unit weight of batteries and opened the modern phase of fuel cell development.

Francis T. Bacon was a remarkable man. Freed from the need for a salary, he could bring to bear his curiosity and his determination full time on the development of one of the principal pathways to a stable world without pollution and planetary warming. After his work had been recognized (a Fellowship in the Royal Society and a substantial prize from NASA), he continued to stand behind world-wide fuel cell development into his 80s. He corresponded with laboratories in many countries, following up this lead, encouraging that demonstration. A discussion with the 80-year-old Bacon was so intense, so fast, so fact filled, that an hour of the vigorous interchange demanded was more than most fuel cell experts 40 years younger could maintain.[1]

13.3. EFFICIENCY

13.3.1. Maximum Intrinsic Efficiency in Electrochemical Conversion of the Energy of a Chemical Reaction to Electric Energy

To obtain maximum efficiency in electrochemical conversion, let an essential thermodynamic equation be recapitulated. It is

$$-\Delta G = W_{rev} - P\Delta V \tag{13.1}$$

---

[1] A study of Bacon's seminal contribution to the development of fuel cells strongly supports the theory of the dominating influence of genes in forming character and ability. Francis Thomas Bacon was a member of the family of Francis Bacon, the great Baron Verulam, who lived in the seventeenth century and is regarded by historians of science to vie with Newton, Leibnitz, and Kepler as among those very few who founded the idea of science as we know it. Tom was descended directly, he told me, from Sir Nicholas Bacon, Lord Keeper of the Great Seal of Elizabeth I. Francis Bacon was a relative of Sir Nicholas.

This means the following: The change in free energy in a reaction is equal to the *total reversible work* obtainable from the reaction (this includes *all* kinds of work, i.e., gravitational, electrical, surface, etc., and also the work of expansion) diminished by the work of expansion, $P\Delta V$. Hence,

$$-\Delta G = W'_{rev} \tag{13.2}$$

where W'_{rev} is all the work obtainable from the reaction exclusive of any work that can be obtained from a possible volume change in the system.

Now, if one carries out a chemical reaction in an electrochemical way, then in example of the reaction $2H_2 + O_2 \rightarrow 2H_2O$, there will be two partial reactions

$$2H_2 \rightarrow 4H^+ + 4e \tag{13.3}$$

and

$$O_2 + 4H^+ + 4e \rightarrow 2H_2O \tag{13.4}$$

After each has been carried through with the stoichiometric quantities (and simultaneously) one has in fact carried out the overall reaction $2H_2 + O_2 \rightarrow 2H_2O$ and hence the normal change of free energy associated with this reaction at a given temperature and pressure has occurred.

However, something else has occurred, too, namely, the transport of four electrical charges across a total potential difference of V, the cell potential. Since thermodynamics calculations are based on the assumption that the chemical change has been carried out near equilibrium (i.e., in a reversible way), the cell potential V to which one refers here is the thermodynamic equilibrium potential $V = V_e$. It is that obtained on an electronic voltmeter with the velocity of the electrode reactions being "infinitely slow" in conformity with the conditions of thermodynamic reversibility. Now, the electrical work of transporting such charges (four electrons per two molecules of water formed, or 4 faradays, if molar quantities are considered) is the total charge transported multiplied by the potential difference through which it passes, i.e., $4FV_e$. Thus, the general expression for the change of free energy in one act of an electrochemical reaction in which the number of electrons transported externally for each act of the equivalent electrode reactions is n, is

$$nFV_e \tag{13.5}$$

Compare this *electrical* work carried out in the reaction with W'_{rev}, the total work obtainable from the reaction, excluding volume-change work. What other kind of work than electrical work is obtainable from the reaction $2H_2 + O_2 \rightarrow 2H_2O$ *carried out in this electrochemical way?* There is no surface work and no gravitational work. Carried out in this electrochemical way and in an ideal manner (i.e., infinitely slowly so that the potential differences in the cell are those characteristic of equilibrium),

TABLE 13.1
Theoretical Cell Potentials of Various Oxidation Reactions at 25 °C

Fuel	Reaction	$\Delta G°$ (kJ mol^{-1})	$V_e°$ (V)
Hydrogen	$H_2 + \frac{1}{2}O_2 \rightarrow H_2O$	−56.69	1.229
	$H_2 + Cl_2 \rightarrow 2HCl$	−62.70	1.370
Propane	$C_3H_3 + 5O_2 \rightarrow 3CO_2 + 4H_2O$	−503.90	1.093
Methane	$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$	−195.50	1.060
Carbon monoxide	$CO + \frac{1}{2}O_2 \rightarrow CO_2$	−61.45	1.333
Ammonia	$NH_3 + \frac{3}{4}O_2 \rightarrow \frac{3}{2}H_2O + \frac{1}{2}N_2$	−80.8	1.170
Methanol	$CH_3OH + \frac{3}{2}O_2 \rightarrow CO_2 + 2H_2O$	−168.95	1.222
Formaldehyde	$CH_2O + O_2 \rightarrow CO_2 + H_2O$	−124.7	1.350
Formic acid	$HCOOH + \frac{1}{2}O_2 \rightarrow CO_2 + H_2O$	−68.2	1.480
Hydrazine	$N_3H_4 + O_2 \rightarrow N_2 + 2H_2O$	−143.9	1.560
Zinc	$Zn + \frac{1}{2}O_2 \rightarrow ZnO$	−76.05	1.650
Sodium	$Na + \frac{1}{2}H_2O + \frac{1}{4}O_2 \rightarrow NaOH$	−71.84	3.120
Carbon	$C + O_2 \rightarrow CO_2$	−94.26	1.020

$$W'_{rev} = nFV_e \tag{13.6}$$

Hence, from Eq. (13.2),

$$-\Delta G = nFV_e \tag{13.7}$$

Table 13.1 lists the Gibbs free-energy change and the corresponding equilibrium–potential differences for the reactions of the oxidation of some currently used and potential fuels.

It is in this sense it is said that in an electrochemical energy converter, the ideal maximum efficiency is 100% for, as in the above idealized situation, if one could carry out reactions in such a way that the electrode potentials were infinitely near the equilibrium values, the electrical energy one could draw[2] from the reaction would be nFV_e and this is all of the free-energy change ΔG, which is the maximum amount of useful work one can obtain from a chemical reaction.

---

[2]One would draw it out by making it drive a current through an external load (e.g., the armature of an electric motor), which ideally would convert the electrical energy to mechanical energy at 100% efficiency and which in fact does carry out such a conversion at more than 95% efficiency.

What has been shown, then, is that the intrinsic maximum efficiency of the electrochemical converter working under ideal conditions is 100% of the ΔG, which is the *useful* or intrinsically *available* work of a chemical reaction. This is an encouraging fundamental result for electrochemical energy conversion. It shows a clear and unique advantage over the efficiency of classical thermal energy converters. It also shows an intrinsic advantage when a comparison is made with other, newer direct energy converters because most of these, such as the thermoelectric device in which heat is taken in at one temperature and rejected at a lower temperature, are also subject to the debilitating Carnot efficiency limitation. Herein lies, then, the unique and attractive potential of the electrochemical method of converting the energy of chemical reactions to energy in the form of electricity.

However, in making a numerical comparison with other types of converters, something has been done here that is not fair. A comparison has been made with the *available* energy in a reaction, ΔG, and it has been shown that the electrochemical method of energy conversion could convert to electricity all the energy intrinsically available as the result of a chemical reaction (independently of the method of conversion). Not quite all the energy difference between the reactants and products of an electrochemical reaction can be made available, however, even by the electrochemical method, because some of it is wasted in very fundamental processes connected with the ordering and disordering (i.e., the entropy losses and gains) that also occur in chemical reactions. It is the enthalpy change (or change in heat content) ΔH that is equivalent to the total change in energy between the reactants and products of a reaction, *including* the energy lost in entropy increases. It is a more significant standard of comparison, therefore, to base the efficiency of any energy-conversion method on a comparison of how much energy it gives compared with a change in heat content (enthalpy), ΔH, in a reaction because ΔH is the total energy difference between the products and reactants of a reaction. The ΔH is usually larger in magnitude than ΔG, often by 10 to 20% (see Table 13.2). Hence, a second and better expression [see also Eq. (13.7)] for the intrinsic maximum efficiency of an ideal electrochemical converter is

$$\varepsilon_{max} = \frac{\Delta G}{\Delta H} = -\frac{nFV_e}{\Delta H} \tag{13.8}$$

It can be seen from Eq. (13.8) that there is no general single number (e.g., 100%) that one can give for the maximum intrinsic efficiency of an electrochemical energy converter on a heat content basis. Examples of values for typical overall reactions that are or might be used in fuel cells are given in Table 13.2. The values can depend on whether the cell reaction is carried out in an acid or alkaline electrolyte since in the latter case, the reaction product is a carbonate with a somewhat different standard free energy than CO_2. Since the activity of the reactants and products depends on the concentration of the electrolyte, so does the potential of the cell reactions and the efficiency of conversion. Still, one may say that the maximum intrinsic efficiency for electrochemical energy conversion based on a comparison of heat content is in the

TABLE 13.2
Standard Free Energy, Enthalpy, and Maximum (Intrinsic) Efficiency for Some Possible Fuel-Cell Reactions

Reaction	$-\Delta G^\circ$ (kcal)	$-\Delta H^\circ$ (kcal)	V_e (V)	Efficiency (%) $(\varepsilon)_{max} = \dfrac{-nFV_e}{\Delta H} \times 100$
$H_2 + \frac{1}{2}O_2 \rightarrow H_2O$	56.69	68.32	1.229	83
$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$	195.50	212.80	1.060	92
$C_3H_3 + 5O_2 \rightarrow 3CO_2 + 4H_2O$	503.90	530.61	1.093	95
$C_{10}H_{22} + \frac{1}{2}O_2 \rightarrow 5CO_2 + 11H_2O$	1574.42	1632.33	1.102	97
$CH_3OH + \frac{3}{2}O_2 \rightarrow CO_2 + 2H_2O$	168.95	182.61	1.222	93
$C + \frac{1}{2}O_2 \rightarrow CO$	32.81	26.42	0.712	124
$CO + \frac{1}{2}O_2 \rightarrow CO_2$	61.45	67.63	1.333	91
$H_2 + Br_2 \rightarrow 2HBr$	24.57	28.90	1.066	85

region of 90%[3] compared with heat engines which, when operating in currently tolerable temperature intervals, have a maximum intrinsic efficiency of 20 to 40%. The higher the T_H in the Carnot expression, the better the efficiency of heat engines. If one raises the T_H to about 1500 °C (!), the Carnot efficiency of heat engines *could* rise toward the normal efficiency of fuel cells at 190 °C.

13.3.2. Actual Efficiency of an Electrochemical Energy Converter

It has been shown (Section 13.3.1), that in an electrochemical energy converter, the maximum cell potential is the value V_e obtainable when the reaction in the cell is electrically balanced out to equilibrium, i.e., when no current is being drawn from the cell. *As soon as the cell drives a current through the external circuit, the cell potential falls from the equilibrium value V_e to V.* The value of the actual potential V at which the cell works when delivering a current i is always *less* than the equilibrium potential V_e. Hence, one has from Eq. (13.8)

---

[3]Note that in the hypothetical converter in which one would realize the reaction $C + 1/2O_2 \rightarrow CO$, the efficiency of conversion on a heat-content basis could be *greater* than 100%. Thus, more electrical energy would be obtained from the system (worked in an ideal reversible way) than the difference in the heat content of the products and reactants. This is because the entropy change is positive in the reaction quoted, i.e., the disorder of the product, 1 mol of gas, is greater than the disorder of the reactants, 1/2 mol of gas. The cell would tend to cool upon working, and heat energy would be absorbed from the surroundings and converted to electricity if it were arranged for the converter to continue to work isothermally.

$$\varepsilon_0 = -\frac{nFV_e}{\Delta H}\frac{V}{V_e} \tag{13.9}$$

or

$$\varepsilon_0 = \varepsilon_{max}\varepsilon_p \tag{13.10}$$

where ε_{max} is the maximum efficiency given by Eq. (13.8), and ε_p is known as the *voltage efficiency* given by

$$\varepsilon_p = \frac{V}{V_e} \tag{13.11}$$

Of course, this picture is true only if the reactants are completely converted to final reaction products, i.e., if the overall reaction is fully accomplished and none of the electrons take part in some alternative reaction. To allow for the possibility that such a wastage does occur, we must consider the current or Coulombic efficiency ε_f to take into account the incomplete conversion of reactants into products. The overall efficiency will be

$$\varepsilon_0 = (\varepsilon_{max}\varepsilon_p)\varepsilon_f \tag{13.12}$$

In many reactions of interest, ε_f is virtually unity.

13.3.3. Physical Interpretation of the Absence of the Carnot Efficiency Factor in Electrochemical Energy Conversion

At first, one might conclude that there is no need for an explanation of why electrochemical energy converters differ both from classical indirect energy converters and also from other direct energy-conversion devices in lacking the intrinsic limitation in the conversion of the energy of a chemical reaction as a consequence of Carnot's theorem. Those with Carnot terms are all *heat engines.* Fuel cells, on the other hand, do not use the heat given out in a reaction but rather separate the reactants (which do not collide and react and expand a gas to do $p\,dV$ work) and make them yield equivalent electric charges, which then create magnetic fields that turn the armatures of motors, etc.

However, there is something so important and fundamental in this different way of converting the energy released in a chemical reaction to work that it is worth a more detailed inquiry on a molecular level, concerning the difference in intrinsic conversion-efficiency maxima between thermal and electrochemical reactions.

In a thermal reaction, the *mechanism* of energy conversion may be visualized on a molecular scale in the way that the change in energy between products and reactants is released in the form of the heat or kinetic energy of the constituent gases. These then impact upon some mechanical and movable object (e.g., the piston in an internal

combustion engine), but in these collisions *there is a series of glancing angles of incidence of the molecules on the cylinder and the transfer of momentum to the piston from the* (hot) *molecules is not complete.* The degree of completeness is measured by the change in the kinetic energy of the particles after they have struck the piston (Fig. 13.4). The temperature of the gases initially—i.e., before striking the piston—is T_1 or the kinetic energy per particle, $\frac{3}{2}kT_1$. Afterward, it is T_2 or the kinetic energy per particle, $\frac{3}{2}kT_2$. the point is that only a part of the kinetic energy, $\frac{3}{2}kT_1 - \frac{3}{2}kT_2$, or heat energy was transferred to the piston; the rest of the energy escaped, i.e., remained as heat in the gas that was rejected to the heat sink (the products of the reaction are emitted from the reaction chamber at a lower temperature than the maximum, but they still contain much of the reaction's energy, i.e., they are still *hot*). The efficiency of the conversion process is given by

Heat source at
temperature T_1

Hot
mole-
cules
with
heat
Q_1

Work W

Cold
mole-
cules
with
heat
Q_2

$$\epsilon = \frac{Q_1 - Q_2}{Q_1} = \frac{W}{Q} = \frac{T_1 - T_2}{T_1}$$

Heat sink at
temperature T_2
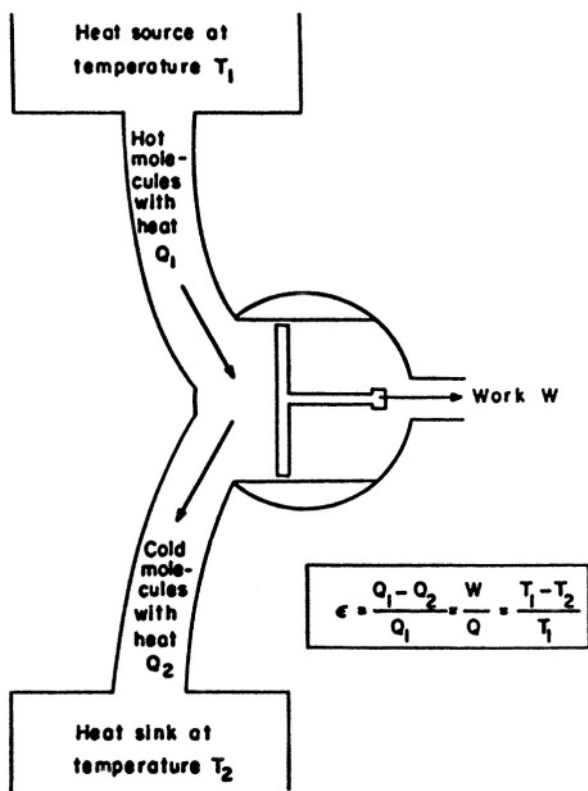
Fig. 13.4. Representation of the efficiency of a heat engine.

$$\varepsilon = \frac{\frac{3}{2}kT_1 - \frac{3}{2}kT_2}{\frac{3}{2}kT_1} = \frac{Q_1 - Q_2}{Q_1} = \frac{W}{Q_1} = \frac{T_1 - T_2}{T_1} \tag{13.13}$$

In the electrochemical converter, there are no collisions between the particles (e.g., H_2 and O_2) that react to form, e.g., H_2O. They simply undergo electron supplying (for the H_2) and electron accepting (for the O_2) reactions on the electrodes. These electrons travel around an external circuit and while doing so pass through a *load* (e.g., the armature of a motor) and thereby do work. They complete the reaction and allow the product water to be formed. Hence, there is no analogue of any process of incomplete transfer of the energy difference between the two sides of a reaction to the outside of the convert, as there is in thermal conversion. The electrons pass through the *entire* potential difference generated by the cell reaction and not through only a part of it.

It may, then, be asked why an electrochemical converter working under ideal conditions does not convert into electric energy *all* the energy released in a chemical reaction, ΔH, but only the free-energy change in the reaction, ΔF. This is of course because the reaction taking place is still the same overall chemical reaction as in the thermal case, i.e., it is still $2H_2 + O \rightarrow 2H_2O$. The *entropy* change will be unaltered. A part of the ΔH is used up in the unavailable energy connected with the differences in order and disorder between the products and the reactants, and this is unavoidably so, independent of any method of energy conversion.

13.3.4. Cold Combustion

These considerations lead to an explanation of why reactions between some substance and oxygen that are carried out in an electrochemical way are called *cold combustion* (Justi and Winsel). The *net* cell reaction (the summation of the two electrode reactions $2H_2 \rightarrow 4H^+ + 4e$ and $O_2 + 4H^+ + 4e \rightarrow 2H_2O$ in which the electrons cancel out) is identical to the actual combustion reaction; it gives out heat that may be converted to mechanical work with an efficiency given by the Carnot expression. The rest of the heat is evolved as heat. But in the electrochemical reaction, the heat ΔH is not given off. What is given off is not hotter molecules (made hotter by transfer of the difference of the potential energy of the reactants and products to kinetic energy), but a stream of electrons, the total energy of which is ΔG (for each act of the overall reaction). The combustion reaction has occurred, but it is *cold*.

It has been said that the same reaction has occurred as in combustion, and the energy ΔG has been electrically drawn off. The total energy change in the reaction, however, is ΔH. Hence, there *is* some heat energy given up or taken in during the ideal reversible working of an electrochemical reactor. It is $\Delta H - \Delta G$ or $T\Delta S$ and is usually negative, i.e., heat is given off. The amounts are small, e.g., in the oxidation of 1 mol of propane into carbon dioxide, $\Delta H = -530.6$ kcal mol^{-1}, but $T\Delta S$ is only -32 kcal mol^{-1}. Electrochemical combustion is *almost* cold.

13.4. KINETICS OF FUEL CELL REACTIONS

13.4.1. Making *V* near *Ve* Is the Central Problem of Electrochemical Energy Conversion

One cannot change I_{max} for a given reaction, and ε_f is usually near unity [see Eq. (13.12)]. Consequently, the main efficiency-determining quantity subject to variation is *V*, the actual cell potential. Thus, the overall efficiency of electrochemical energy converters depends on how the overall cell potential varies with the current density the cell is producing. One might say that it depends on how much of the energy of its own reaction the cell has to use up to et its two electrode reactions to take place at the desired rate, the rest of the energy being available for use outside the cell, i.e., for useful work.

In considering the behavior of electrochemical systems in action when a current is flowing through them, an expression was developed earlier for the cell potential *V* as a function of the overall current *I*. It was shown that for a very simple electrochemical energy converter having electrodes of the same area *A* and delivering a current *I*, one has

$$V = IR_e = (E_{e,\mathrm{SO}} - \eta_{an,\mathrm{SO}} - \eta_{cath,\mathrm{SO}}) - (E_{e,\mathrm{SI}} + \eta_{an,\mathrm{SI}} + \eta_{cath,\mathrm{SI}}) - IR_i$$

$$= (E_{e,\mathrm{SO}} - E_{e,\mathrm{SI}}) - \eta_{an,\mathrm{SO}} - \eta_{cath,\mathrm{SO}} - \eta_{an,\mathrm{SI}} - \eta_{cath,\mathrm{SI}} - IR_i \quad (13.14)$$

where *E* is an equilibrium potential, η is overpotential of the type and location indicated, R_e is the resistance external to the cell, and R_i is the resistance of the space between the electrodes. If both the deelectronation and the electronation reactions are running under high-field, or Tafel, conditions, the high-field approximation can be used to relate the two activation overpotentials (that for the source electrode and that for the sink electrode) to the current density

$$\eta_{an,\mathrm{SO}} = \frac{RT}{\alpha_{\mathrm{SO}}F} \ln \frac{(I/A_{\mathrm{SO}})}{i_{0,\mathrm{SO}}} \quad (13.15)$$

and

$$\eta_{an,\mathrm{SI}} = \frac{RT}{\alpha_{\mathrm{SI}}F} \ln \frac{(I/A_{\mathrm{SI}})}{i_{0,\mathrm{SI}}} \quad (13.16)$$

Further, the expression for the concentration overpotential can be substituted for $\eta_{cath,\mathrm{SO}}$ and $\eta_{cath,\mathrm{SI}}$. They are

$$\eta_{cath,\mathrm{SO}} = \frac{RT}{nF} \ln \left(1 - \frac{I/A_{\mathrm{SO}}}{i_{L,\mathrm{SO}}} \right) \quad (13.17)$$

and

$$\eta_{\text{cath,SI}} = \frac{RT}{nF} \ln\left(1 - \frac{I/A_{\text{SI}}}{i_{L,\text{SI}}}\right)$$

(13.18)

Thus, the expression for the cell potential becomes

$$V = IR_e = V_e - \left[\frac{RT}{\alpha_{\text{SO}}F} \ln \frac{I/A_{\text{SO}}}{i_{0,\text{SO}}} + \frac{RT}{nF} \ln\left(1 - \frac{I/A_{\text{SO}}}{i_{L,\text{SO}}}\right)\right]$$

$$- \left[\frac{RT}{\alpha_{\text{SI}}F} \ln \frac{I/A_{\text{SI}}}{i_{0,\text{SI}}} + \frac{RT}{nF} \ln\left(1 - \frac{I/A_{\text{SI}}}{i_{L,\text{SI}}}\right)\right] - IR_i$$

(13.19)

Equation (13.19) represents a cell–potential–current relation (Fig. 13.5) over a wide range of conditions; that is why it appears complicated even though it is given here for the idealized case of a converter with two planar, smooth electrodes so that
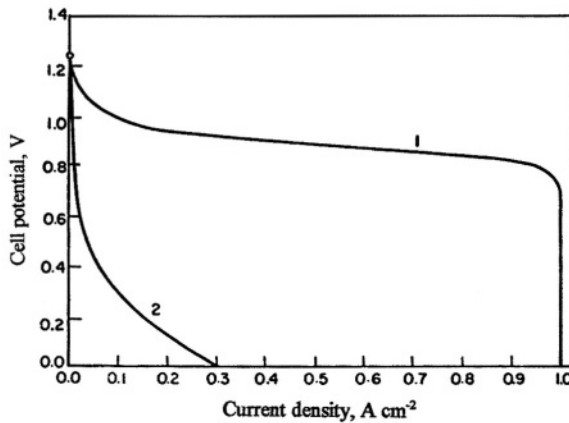
Fig. 13.5. Cell-potential vs. current-density relations for an idealized electrochemical energy converter with planar, smooth electrodes. Assumed are the following parameters: Curve 1, $i_{0,\text{SI}} > 1$ A cm^{-2}, $i_{0,\text{SO}} = 10^{-3}$ A cm^{-2}; $i_{L,\text{SI}} = 1$ A cm^{-2}, $i_{L,\text{SO}} = 1$ A cm^{-2}, $\alpha_{\text{SI}} = \infty$, $\alpha_{\text{SO}} = 1/2$; $R_i = 10^{-2}$ ohm. Curve 2, $i_{0,\text{SO}} = 10^{-6}$ A cm^{-2}; $i_{L,\text{SI}} = 1$ A cm^{-2}, $i_{L,\text{SO}} = 1$ A cm^{-2}; $\alpha_{\text{SI}} = 1/2$, $\alpha_{\text{SO}} = 1/2$; $R_i = 1$ ohm.

the complexities of the cell potential vs. current relation with porous electrodes is avoided.

At a current density (Fig. 13.6) sufficiently far below the limiting current density values [see Eq. (13.19)] and when the ohmic losses inside the cell are negligible, the activation-overpotential terms dominate the expression for the relation of current to potential, i.e.,

$$V \cong V_e - \frac{RT}{\alpha_{SO}F} \ln \frac{I/A_{SO}}{i_{0,SO}} - \frac{RT}{\alpha_{SI}F} \ln \frac{I/A_{SI}}{i_{0,SI}} \qquad (13.20)$$

At higher current densities than those referred to in Eq. (13.20), the activation-overpotential terms in this equation change much less with current than the ohmic overpotential owing to the internal resistance of the cell. Under these conditions, when $(I/A)/i_L$ continues to remain negligible and the variation of V and I (but not its absolute value) is dominated by the IR_i term, one has (Fig. 13.7)

$$V \cong V_3 - \text{const} - IR_i \qquad (13.21)$$
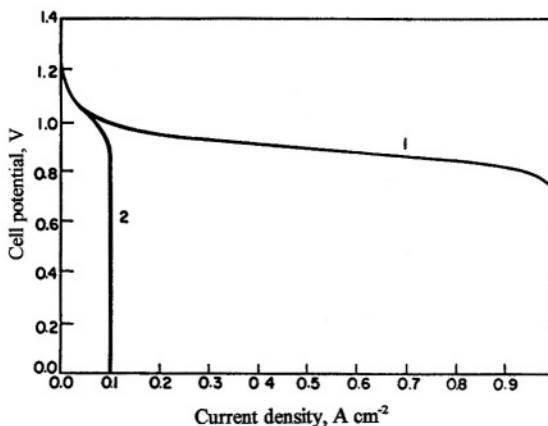
Fig. 13.6. Cell potential vs. current density relations for an electrochemical energy converter for negligible values of IR, showing the influence of a limiting current. Curve 1, kinetic parameters correspond to curve 1 in Fig. 13.5; curve 2, the same as curve 1 except $i_{L,SI} = 10^{-1}$ A cm^{-2} and $i_{L,SO} = 10^{-1}$ A cm^{-2}.
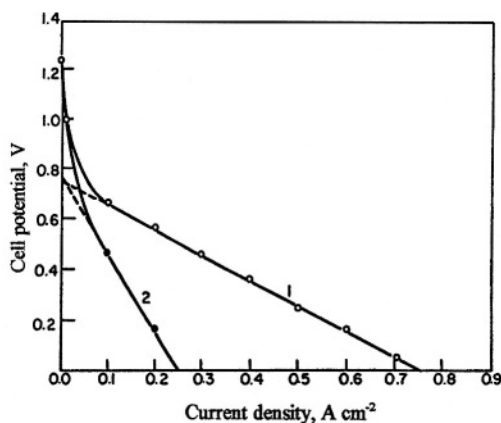
Fig. 13.7. Graphical representation of the influence of the internal resistance of an electrochemical energy converter on the cell potential when mass-transfer polarization is negligible. The early nonlinear part of the curve represents the effect of the activation overpotential on the cell potential before ohmic polarization has become important.

where the constant represents the activation overpotential, which changes more slowly with current density than does the linear ohmic term.[4] At sufficiently high current densities, the I/A of Eq. (13.19) starts becoming comparable with i_L and concentration overpotential starts to reduce the cell potential in a more significant way than the IR_i term, which may now be taken as relatively constant. Thus (Fig. 13.8),

$$V \cong V_e - \text{const 2} - \frac{RT}{nF} \ln\left(1 - \frac{I/A_{SO}}{i_{L,SO}}\right) - \frac{RT}{nF} \ln\left(1 - \frac{I/A_{SI}}{i_{L,SI}}\right) \qquad (13.22)$$

It is seen, therefore, that the cell potential V and consequently the efficiency of an electrochemical converter (Fig. 13.9) are determined by the activation overpotential, by the electrolyte conductance, and by mass transfer (i.e., the solubility of the reactants). The factors that dominate the way the efficiency of the conversion of energy *changes* with an increase in current density are at low current density, the activation

---

[4]Note how, even in the region in which there is linear behavior of V with respect to I, the *actual value* of the potential that the generator could put out depends on the value of the so-called "constant," i.e., on the activation overpotential and thus on the exchange current densities and the catalytic power of the electrodes.
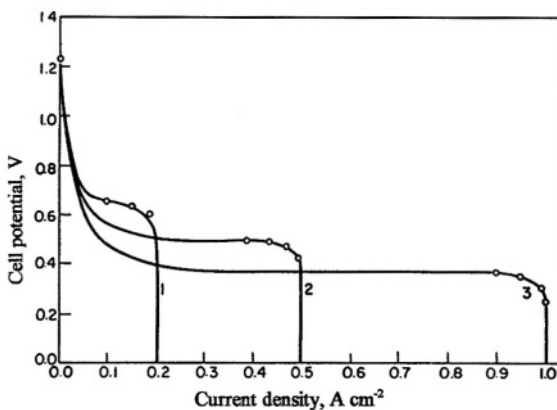
Fig. 13.8. Cell potential vs. current density relations for an electrochemical energy converter (curve 1 of Fig. 13.5), showing the influence of mass-transfer limitations. Curve 1 is for $i_L = 0.2$ A cm^{-2}; curve 2, for $i_L = 0.5$ A cm^{-2}; and curve 3, for $i_L = 1.0$ A cm^{-2}.

overpotential; at medium current density, the electrolyte resistance; and, at the highest current densities, the mass transport. These factors are the ones that *dominate* at a given condition in causing changes in the efficiency (or power) in that particular region. But the absolute value of quantities describing the cell potential versus current density behavior is determined by the sum of the influence of the activation, ohmic, and diffusional overpotentials. Thus, when mass transport becomes the most important factor in causing the efficiency to change, the actual efficiency of conversion at the point at which diffusion becomes important is also influenced by the value of the activation overpotential, which dominated changes in the situation earlier, and by ohmic influences, which are the main causes of increase in overpotential and decrease in cell potential in the middle range of current density (Fig. 13.9).

13.4.2. Electrochemical Parameters That Must Be Optimized for Good Energy Conversion

Equation (13.19) also makes another aspect clear, the parameters that one has to attempt to optimize in electrochemical energy converters. Inspection of Eq. (13.19) and Figs. 13.5 and 13.8 shows that ideal reversible behavior will be approached when i_0 and i_L are very large and the internal electrolyte resistance of the cell is small. The maximization of i_L is a matter of designing and engineering cells to which the diffusion and convection of ions most easily occurs (see Section 13.5.1). To reduce R_i, electro-
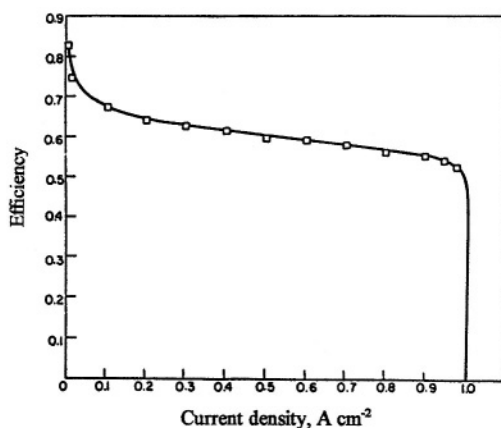
Fig. 13.9. Efficiency vs. current density rela-
tions for an electrochemical energy converter
(calculated for curve 1 of Fig. 13.5).

lytes of as high a conductance as possible are used. The conductance of $1-5\,M\,H_3SO_4$ and KOH is in the range of $0.5\,\text{ohm}^{-1}\,\text{cm}^{-1}$, while that of most other concentrated aqueous solutions is much smaller. Minimizing R_i is also largely a matter of structural arrangement of the electrodes in the cell, in particular with respect to the type of electrodes (usually *porous*) used.

A great deal of the future of electrochemical energy conversion depends upon a detailed understanding of two electrode kinetic parameters—the exchange-current density (i_0) and Tafel slope, b. It is through these parameters that electrochemical energy conversion becomes linked up with electrocatalysis (see Sec. 7.2.1). The values of i_0 observed for some reactions vary over many orders of magnitude (e.g., $\sim 10^{-7}\,A\,\text{cm}^{-2}$ for methanol as a fuel on Pt (see Table 13.3) at different catalysts, and the aim of fundamental research on electrocatalysis is to understand the phenomenon of this considerable dependence on the electronic structure of the substrate. This understanding may make it possible to make electrodes with high i_0 values (and small polarization) at high rates for fuels that are cheap and that have a large amount of energy per gram and per dollar available for fuel cell use. Only the relative order of catalytic activity among various substrates is known for most fuels (see Table 13.4). Regarding the second parameter, b, it is most beneficial to have as low a value as possible to decrease the activation overpotential. As is well known to electrochemists, b is generally equal to 0.12 V/dec if the first electron transfer step in a reaction is rate determining. It decreases to about half its value when the second electron transfer step controls the rate. In practice, the maximum energy-conversion efficiencies obtained in some electrochemical converters are as much as 65%.

TABLE 13.3
Exchange Current Densities and Tafel Slopes for the Oxidation of Methanol to Carbon Dioxide on Various Catalysts

Catalyst composition	$-\log i_0$	Tafel slope (b)
Pt-Fe	4.0	0.103
Pt-Ni	4.3	0.109
Pt-Rh-Fe	5.4	0.081
Pt-Co	5.5	0.084
Pt-Pb	6.0	0.085
Platinum	7.3	0.066
Pt-Rh	7.9	0.063
Pt-Cu	8.4	0.059
Pt-Au	8.8	0.058
Pt-Co	9.4	0.058
Pt-Ni	9.1	0.060
Pt-Fe	6.5	0.073
Iridium	9.9	0.053

13.4.3. The Power Output of an Electrochemical Energy Converter

When an electrochemical converter is working as an engine, i.e., when it is being used to produce power—energy production at a given rate—the efficiency of the conversion of chemical energy to work will decrease with an increase in the power density. For practical applications (e.g., in driving trucks), it is the power output or the rate at which they are able to do work that has to be examined. The power P of an electrochemical energy converter is defined thus

$$P = IV \tag{13.23}$$

from which it follows that the power is small when I is small, even though V is near the maximum V_a. However, the power output is also small when I is very large because the sudden growth of concentration overpotential when the current density approaches i_L tends to drive V to zero [see Eq. (13.19)]. Thus, the P vs. I curve should pass through a maximum (Figs. 13.9 and 13.10).

The distinction between the situation in which one needs predominantly high efficiency and one in which one wants high power becomes clear when one compares the P vs. I and ε vs. I curves (Fig. 13.11). When its efficiency is at a maximum, the electrochemical energy converter is a less good power source. As the current density is increased, the power output increases, but the efficiency decreases. Of course, at the highest current drains, both the power and efficiency fall toward zero.

TABLE 13.4
Relative Efficiencies of Metal Catalysts for Three Reactions

Electrolyte	Catalyst Surface	$CH_3OH \rightarrow CO_2$	$HCHO \rightarrow CO_2$	$HCOOH \rightarrow CO_2$
5 M KOH	Smooth	Pt > Pd, Ir > Au > Rh		
	Rough	Pd, Pt, Ir > Au > Rh		
1 M Na$_2$SO$_4$	Smooth	Ir > Ag, Pt, Pd > Au	Ir > Rh > Pt > Pd > Au	Pd > Ir > Pt, Rh > Au > Ag
	Rough	Au > Pt > Ir > Pd > Rh	Au > Ir, Rh > Pt > Pd	Pd, Ir > Pt, Au > Rh
3.3 M KHCO$_3$	Smooth	Pt > Ir > Pd > Au > Ag > Rh	Ag > Rh, Pt, Ir > Au > Pd	Ag > Pt > Pd > Rh, Ir > Au
	Rough	Pt > Ir > Pd > Au > Rh	Au > Ir > Pt, Rh > Pd	Ir > Rh > Pt > Au
1 M H$_2$SO$_4$	Smooth	Pt, Ir > Pd	Pt, Ir > Pd > Au	Pt > Ir > Pd > Au, Ag
	Rough	Pt, Ir > Pd, Au > Rh		Ir, Pt > Au > Pd > Rh

Source: Reprinted from K. R. Williams. *An Introduction to Fuel Cells*, copyright 1966 with permission from Elsevier Science.
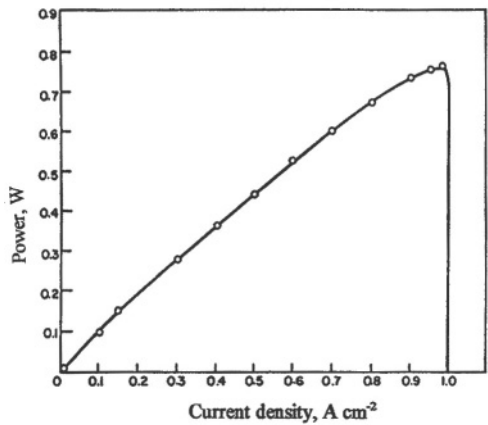
Fig. 13.10. Power vs. current density relations for an electrochemical energy converter (calculated for curve 1 of Fig. 13.5).

13.4.4. The Electrochemical Engine

The principal use of the electrochemical energy converter to date has been a situation (auxiliary power in space) in which the *weight* of the energy converter plus the fuel carried is of primary importance. Thus, in any energy-conversion situation
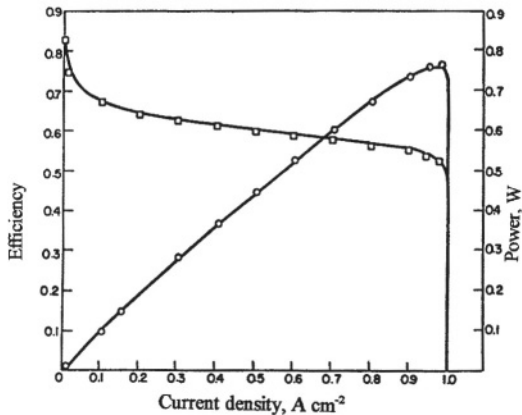
Fig. 13.11. Power vs. current density and efficiency vs. current density relations for an electrochemical energy converter (calculated for curve 1 of Fig. 13.5).

(e.g., transportation in its most general sense) where the system must carry its own fuel to provide a certain number of kilowatt hours of energy during a journey of known duration, the main point is the *efficiency* of conversion; the weight of the fuel necessary for a given operation is clearly inversely proportional to it. Hence the use of fuel cells in space exploited the avoidance of the Carnot cycle efficiency loss in electrochemical conversion.

Another probable use of the electrochemical reactor, however, is as a *power* (rather than energy) source and, in combination with an electric motor, as an *electro-chemical engine* (Henderson). Here, the stress is on power, the rate of delivery of energy per unit time, as well as the efficiency of the conversion.

13.4.5. Electrodes Burning Oxygen from Air

Different materials can be used as oxidants in fuel cells. Yet it is desirable, if possible, to use O_2 from air at one electrode in every earthbound fuel cell because this avoids the necessity of carrying a second fuel for the cathodic reaction. Hence, the cathodic reduction of O_2 has a special importance in electrochemical reactors. The overall reaction in acid solution (see Sec. 7.10) is

$$O_2 + 4H_3O^+ + 4e \rightarrow 6H_2O \qquad (13.4)$$

and in alkaline solution,

$$O_2 + 2H_2O + 4e \rightarrow 4OH^- \qquad (13.24)$$
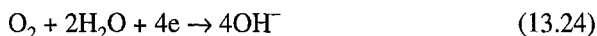
There is a grave disadvantage in this important (and inevitable) electrode reaction. It has an i_0 value in the region of 10^{-10} A cm^{-2}, and hence (from the equation $\eta = RT/\alpha F \ln i/i_0$) the reaction usually contributes considerably to the overpotential in the functioning of an air-burning electrochemical converter. Electrocatalysis of this reaction is needed more than any other in electrochemical energy converters.

13.5 POROUS ELECTRODE

13.5.1. Special Configurations of Electrodes in Electrochemical Energy Converters

In the foregoing material on electrochemical energy converters,[5] the explicit assumption is that the electronic conductors with which the molecules of the fuels

---

[5]The term electrochemical energy converter (eec) is a general one referring to cells in which the introduction of two substances on the electrodes generates power and synthesizes compounds. The eec may thus produce electricity at high efficiency; it may be a power generator in which the construction is mainly aimed at maximizing the power-to-weight ratio; or it may be an electrochemical engine in which the electrochemical generator is connected with an electric motor. *Fuel cell* is the historical name for such devices.
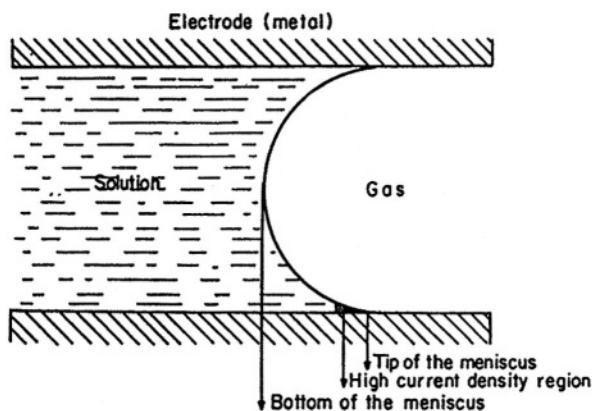
Fig. 13.12. Schematic representation of the three-phase (metal, electrolyte, and gas) interphase in a single pore when the contact angle is a few degrees.

transfer and receive electrons are in the form of planar bodies, e.g., *sheet* electrodes. The presentation is made in this way because it is possible thereby to present simple equations and make relatively clear deductions about the trends arising from changes in the exchange current, conductance, limiting current, and internal resistance.

However, if only planar electrodes were used, much smaller currents would be observed per geometric (or external) unit area of electrode material than are in fact obtained. Electrochemical converters would in fact have no practical uses. Thus, in all actual electrochemical converters, the electrodes are three-dimensional, *porous* structures (e.g., of graphite), whose pores usually contain the *catalyst material* (e.g., platinum) to and from which electric charge transfer occurs.

Earlier workers in this field assumed that the higher currents obtained by using porous electrodes occurred because the electrolyte was in contact with a larger real area of the electrodic catalyst per geometric area than with planar electrodes. However, this view is now regarded as much less than the whole story. The expression for the power of a generator involves the current, and it is desirable to work near the maximum value of this, i.e., near the limiting current (see Sec. 13.4.2). Upon what does this limiting current depend? If it were at a planar electrode, the principal variable would be the diffusion-layer thickness, and this could be altered by various forms of agitation (but the agitation itself uses up some of the power produced and could hardly make the diffusion layer thinner than $\sim 10^{-2}$ to 10^{-3} cm). Consider, however, a three-phase boundary[6] in a porous electrode at which there is a gas in contact with a solution and a metal, as shown in Fig. 13.12. The thickness of such menisci vanishes at the tip and

---

[6] "Three-phase boundary" refers to the unique situation in a meniscus in a fuel cell at which there is a solid (the electrode itself), a liquid (the solution), and a gas (the fuel or air-oxygen), all in contact.
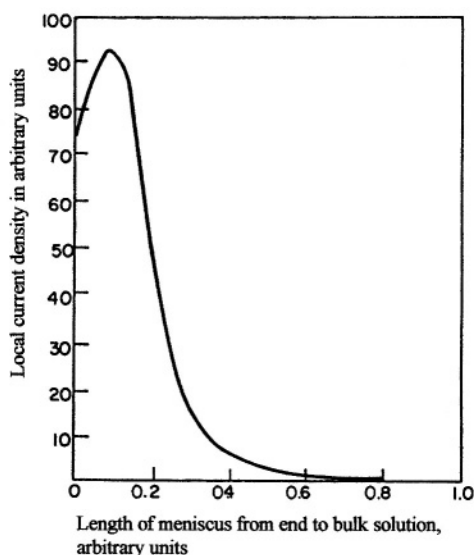
Fig. 13.13. Distribution of current density along the meniscus in a single pore.

increases back toward the bulk of the electrolyte (Srinivasan and Hurwitz, 1967; Cahan, 1969).

These variations in meniscus thickness with distance in a pore imply that there is a very large local decrease in δ in the meniscus (of a *wetted* pore) to less than 10^{-6} cm (Fig. 13.12) and hence a large increase in the limiting diffusion current $[i_L = (D_2 F / \delta) c_{O_2}]$. However, if the thickness of the meniscus is too low, the H_3O^+ ions diffusing back to the bulk of the electrolyte from the dissolution of, say, $H_2 + 2H_2O \rightarrow 2H_3O^+ + 2e$, will not be able to escape sufficiently quickly because of the high resistance (caused by the low cross section) to the diffusion of ions arising from the very thin section implied by such thin menisci. On the other hand, back where the meniscus is sufficiently thick, say, at 0.01 cm from its beginning, δ is only about the same as for diffusion to a planar surface from dissolved gas in a bulk electrolyte. Hence, the highest currents will not be reached at those parts of a pore where a dissolving gaseous fuel has to diffuse through a thick electrolyte or at the very thinnest end of the meniscus wedge, but at some intermediate distance up the meniscus (Fig. 13.13).

The main reason a porous gas electrode is so active,[7] therefore, is that it allows particularly large maximum *diffusion* currents by diffusion through (fairly) thin meniscus layers. But this thesis brings a corresponding antithesis because it implies that farther up the pore where there is no meniscus but bulk solution, the gaseous

---

[7]It is not the large increased *area* provided by each pore, but the high i_L at some small regions of each pore.

reactant dissolved in it is reacting at a rate much less per unit length of meniscus than that in the (moderately thin) meniscus region. Thus, there is little activity in the interior of electrodes of electrochemical reactors; i.e., if the catalyst material present is distributed uniformly throughout the electrode (as in many reactors), much is wasted.

Theory would indicate, therefore, that electrodes should indeed have pores (and hence large numbers of menisci), but they should be very thin or else the catalyst should be concentrated only in a small region of high current density near the higher part of the three-phase boundary. The practical attainment of such a model was first reached in research carried out at the university level (Texas A&M), and in a government research institute (Los Alamos National Laboratory).

13.6. TYPES OF FUEL CELLS

13.6.1. What Is Known So Far about Fuel Cells—Electrochemical Energy Converters

It may be helpful to summarize here what has transpired so far in the chapter:

1. Fuel cells have had an abnormally long development time. The year 1839 was the date of Sir William Grove's discovery. According to a report prepared by the National Science Foundation, the average time for bringing a scientific concept from its first published expression to commercialization is 75 years. Fuel cells have taken more than twice as long.

2. Some see the fuel cell as an electrolyzer worked in reverse. It might be truer to see the electrolyzer as the fuel cell working in reverse. Thus, fuel cells, electro-chemical energy converters, work spontaneously; they go *down* a free energy gradient. To electrolyze water, one has to do something extraordinary—force a reaction to function against its free energy gradient. In any case, the fuel cell and the electrolyzer are in principle the same device; one works in the opposite direction to the other. The fuel cell is a spontaneous energy converter; the electrolyzer is a forced chemical producer.

3. There are two central electrochemical quantities that control the events in a fuel cell. One of them controls the all-important efficiency of conversion. This is the exchange current density (or rate constant) of the electrode reaction at one of the electrodes, usually the cathode. A high i_0 (10^{-3} A cm^{-2}, for example, observed at a fuel cell anode) means a low overpotential and therefore a high efficiency of chemical energy conversion at reasonable power levels (rates of working). A very low i_0 (10^{-10} A cm^{-2}) will mean much energy lost to overpotential and therefore a low efficiency, eventually so low that the principal advantage of the electrochemical compared with the chemical means of energy conversion (the higher efficiency of the former) may be lost (see Fig. 13.14).

4. The second of the two main quantities that control the operation of fuel cells is the limiting current. Were one to use flat plate-type electrodes in fuel cells, the
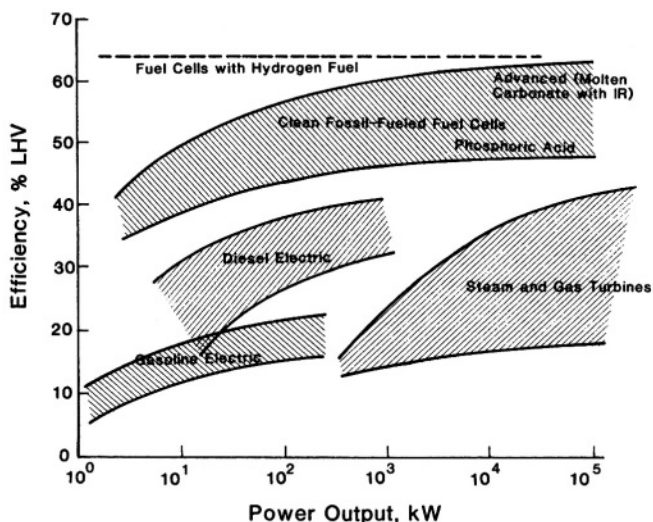
Fig. 13.14. The energy conversion efficiency of different technologies as a function of scale. PAFC, phosphoric acid fuel cell. (Reprinted from *Assessment of Research Needs for Advanced Fuel Cells*, S. S. Penner, ed., Pergamon Press, 1986, pp. 14, 87.)

maximum power density might be less than 1 mw cm^{-2}. Porous electrodes make the limiting current 100–1000 times greater by using the properties of the three-phase boundary, in which there are sections containing very thin layers of solution ($\delta < 0.001$ cm).

5. So, a hypothetical super fuel cell would have $i_0 \gg 10^{-3}$ A cm^{-2}. The efficiency of energy conversion at practical power densities could be greater than 90%, and limiting currents 10^4 times the limiting current densities at planar electrodes giving power densities up to 1 W cm^{-2}. Let us see to what degree real fuel cells approach these hypothetical ideals.

13.6.2. General Aspects of the Practical Fuel Cells

13.6.2.1. *The Cells.*   There are four types of fuel cells: alkaline, phosphoric acid, polymer, high-temperature (molten salt) and finally, the solid (conducting) oxide fuel cell.

Alkaline Fuel Cell (AFC). This cell follows directly from the one that Bacon and Watson produced at Cambridge in the 1950s and is the basis of cells developed for NASA (by International Fuel Cells and predecessor companies (United Technologies Power Systems Divisions, Pratt and Whitney Aircraft) since the Apollo moon program, where pure H_2 fuel is available.

It has a number of advantages over other cells because it functions at relatively low temperatures and its alkaline solutions allow a greater variety of catalysts to be used in it than in the acidic or high-temperature cells. However, the cell has a disadvantage: Air (from which the necessary oxygen must be obtained) contains CO_2. This has to be removed before the air reaches the cell because otherwise the carbonate forced in the alkaline solution blocks pores in the porous electrodes.

Phosphoric Acid Fuel Cell. This cell demonstrates the possibilities of using very concentrated phosphoric acid to allow the temperature of the solution to be raised to about 200–205 °C without the need for high-pressure equipment. The higher temperature makes it possible to produce large amounts of "free" steam for the re-forming of natural gas to the hydrogen upon which most development has been based.

High-Temperature Fuel Cell. There is a cell developed to work with molten carbonates at about 650 °C and a corresponding cell involving a solid oxide electrolyte (yttria-zirconia) having high O-mobility and conductance, and operative at 1000 °C.

Solid Polymer Electrolyte Fuel Cell.[8] Here, there is no apparent liquid solution, or high-temperature ionic conductor. The usual ionic solution between the electrodes is replaced by a well-humidified membrane made of a perfluorosulfonic acid polymer that conducts protons.

Each of these cells has its own place. The alkaline cell is the cell of choice if (as in a space vehicle) supplies of pure hydrogen and oxygen are available, without the H_2 having to be obtained by re-forming from natural gas (whereupon the H_2 would be mixed with CO_2).[9] The phosphoric acid cell works at temperatures > 200 °C and the heat available drives a re-former that produces H_2 from, e.g., natural gas. It avoids the need for pure H_2. The high-temperature cells are envisaged largely for stationary power plants. The high temperature increases i_0, $\eta = RT/\alpha F \ln i/i_0$, and reduces the overpotential needed for a given current density; hence [see Eqs. (13.9) and (13.14)] it increases the efficiency of conversion up toward 70%.

The solid polymer electrolyte fuel cell is that on which the most development work was done in the 1990s because of its projected use in the development of an electrochemical engine for cars. The absence of a bulk liquid component while keeping to temperatures of 80 °C if pure H_2 or H_2 produced from methanol or gasoline on board a vehicle is available, signifies a great advantage. Conversely, the acid environment needs Pt.

13.6.2.2. *Efficiency of Energy Conversion and the Tafel Equation.*

Of the various advantages of direct conversion of chemical fuels to electricity in fuel cells, the greatest is the enhanced efficiency of the electrochemical path to energy. Figure 13.11 is a schematic (based upon deductions made with greatly simplified flat plate cells) of the efficiency of the fuel cell as a function of overpotential. As the rate of

---

[8]More commonly referred to as proton exchange membrane fuel cells since the late 1980s.

[9]The removal of the CO_2 is entirely possible. European work has stressed the advantages of the alkaline fuel cell (e.g., it starts from the cold) more than U.S. work has.

working (power density) increases, the efficiency of energy conversion decreases. For this reason, the b value of the Tafel equation (Sec. 7.2.3) and hence the mechanism of the electrode reaction that fixes the b value, is important [it controls the increase in overpotential (loss of efficiency) as the power level increases]. A cell in which the rate-determining electrode has a b of $2RT/3F$ will exhibit an increase in overpotential (hence, a decrease in efficiency) with an increase of power density at a rate three times *less* than an electrode in which the Tafel slope is $2RT/F$. Hence, the importance of the mechanism of O_2 reduction for the value of the b coefficient in Tafel's equation is dependent on the mechanism of the electrode reactions at which electricity is formed from chemical reactions.

These remarks on the importance of Tafel's b constant are important (and not always brought out very clearly). However, the overriding importance of i_0, exchange current density, in determining efficiency must not be forgotten.

In practice, in all fuel cells that involve the utilization of O_2 from air (Section 13.4.5), the oxygen reduction reaction [Eqs. (13.4) and (13.24)] is always rate determining for terrestrial applications. One can see just how important it is to attempt to develop electrocatalysts for the cathodes of fuel cells on which the enhanced current density is high and Tafel slope is low and the efficiency of energy conversion, therefore, maximal. The direct relation of the mechanism of oxygen reduction and the associated Tafel parameters to the economics of electricity production and transportation is thus clearly seen.

13.6.3. Alkaline Fuel Cells

This cell works optimally at 80 °C using relatively inexpensive materials. When it is switched on in the cold, it produces about one-quarter of the power finally produced after it warms up. This is an advantage compared with other types of fuel cells operating at intermediate (200 °C) or high (650 to 1000 °C) temperatures, which need an auxiliary power source to start them and warm them up. The alkaline environment means that a wide range of electrode catalysts are available, while cells using acid solutions can only use noble metal electrode materials, which is a distinct economic disadvantage for terrestrial applications.

The basic design of one unit cell of the modified Bacon cell used in the Apollo moon project is shown in Fig. 13.15. It operated at about 260 °C at 4 atm pressure on pure hydrogen and oxygen. The high operating temperature meant that no wetproofing agent (e.g., Teflon) could be used in the electrodes, which had a dual-pore structure (Fig. 13.16). A dual-pore structure shows that the electrolyte invaded the small pores as a result of capillary action. The gases are in the big pores and meet the liquid and dissolve in it at the end of the capillaries. The large improvement in cell performance (i.e., high potential at larger current densities) during the first decade since the unveiling of Bacon's 5-kW cell in 1959 is shown in Fig. 13.17.

Historically, and in space, alkaline fuel cells are run on pure H_2 and O_2. There is as yet no massive terrestrial source of hydrogen (e.g., from photovoltaics and the
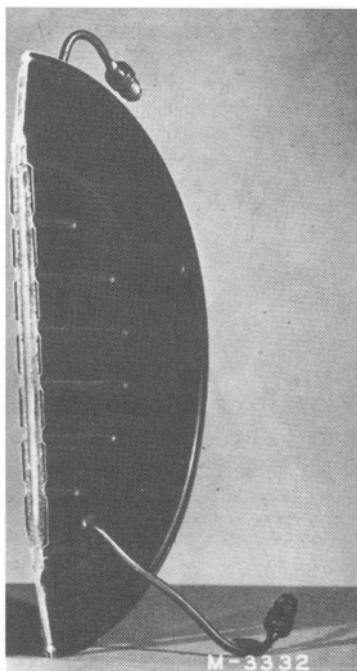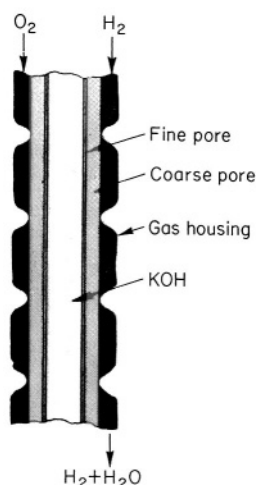
Fig. 13.15. The sandwich design for a single cell in the Pratt and Whitney hydrogen-oxygen fuel cell (1965–1975, used in the Apollo project). (Reprinted from J. O'M. Bockris and S. Srinivasan, *Fuel Cells: Their Electrochemistry*, Fig. 23, copyright 1969. Reproduced with permission of the McGraw-Hill Companies.)

electrolysis of water), so that the economic way to get H_2 is by re-forming natural gas. But 80 °C is not hot enough to raise steam for the re-forming reaction and hence some of the fuel has to be used to produce steam for the re-former. Calculations show that even then, the alkaline cell may show slightly better economics than that using phosphoric acid (Section 13.6.4). On the other hand, the need to remove CO_2 from the airborne O_2 supply is an inconvenience.

13.6.4. Phosphoric Acid Fuel Cells

A combination of 98% H_3PO_4 and 2% water provides a liquid that can be heated to > 200 °C at atmospheric pressures. A high temperature of 150 °C is required to polymerize phosphoric acid to pyrophosporic acid ($H_4P_2O_7$), which has a considerably higher ionic conductivity than the parent acid. It was necessary to raise the operating temperature of the fuel cell to 200 °C in order to tolerate a carbon monoxide level of

Cross section of sintered electrode 25OX
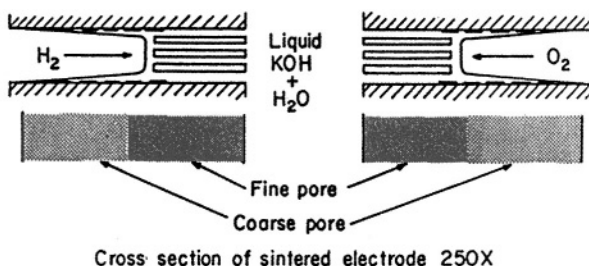
Fig. 13.16. Structure of porous electrode in the Bacon-Pratt and Whitney fuel cell. (Reprinted from J. O'M. Bockris and S. Srinivasan, *Fuel Cells: Their Electrochemistry*, Fig. 22, copyright 1969. Reproduced with permission of the McGraw-Hill Companies.)
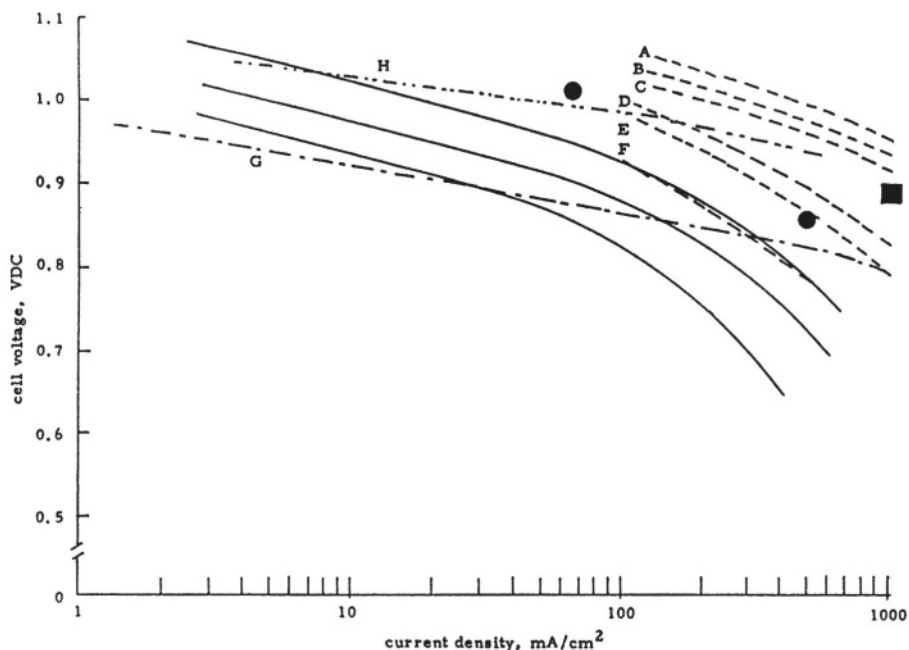
Fig. 13.17. Performance of advanced lightweight pressurized alkaline fuel cells. The dashed lines show initial advanced AFC cell results. A, 149 °C, 17 bar; B, 140 °C, 17 bar; C, 127 °C, 17 bar; D, 110 °C, 4 bar; E, 82°C, 4 bar; F, 82 °C, 1 bar; G, 0.2 MgPt-C and the same conditions as F (IR-free); H, 10 mg/cm^2 Au/Pt, 127 °C, 1 bar (IR-free). •, nominal performance of space shuttle cell (1000 h); ■, United Technologies target goal (1000 hr). Solid lines show solid polymer electrolyte cells for comparison under different pressure and temperature conditions. (Reprinted from *Assessment of Research Needs for Advanced Fuel Cells*, S. S. Penner, ed., Pergamon Press, 1986, pp. 14, 87.)
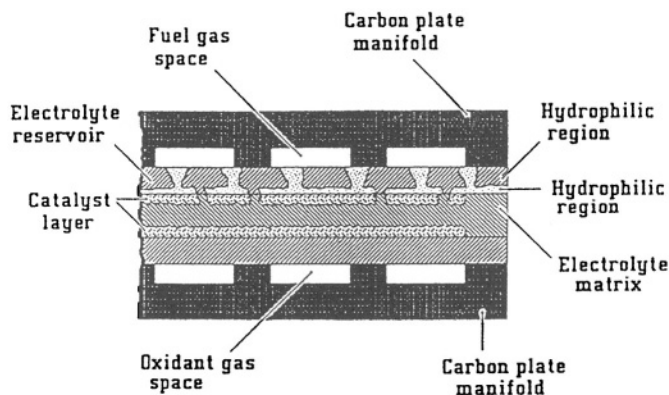
Fig. 13.18. Sandwich configuration of phosphoric acid fuel cells developed by United Technologies. (Reprinted from J. O'M. Bockris and S. Srinivasan, *Fuel Cells*, p. 179, copyright 1993, McGraw-Hill. Reproduced with permission of The McGraw-Hill Companies.)

1–2% in the hydrogen produced by the re-forming reaction for the anodic reaction in the fuel cell. This was seen in the 1960s as a brute-force method to achieve better performance with less reliance on the electrocatalyst which, nevertheless, is used, being generally Pt related (e.g., Pt-Co-Cr alloys) at the cathode. The high temperature gives the heat for re-forming CH_4 to provide H_2, which is always the real fuel (because of its high i_0 value in the reaction of the dissolution of hydrogen at the anode, hence a small overpotential). The extremely concentrated H_3PO_4 has the disadvantage of freezing when the cell is turned off, so its temperature must be maintained above 40 °C by auxiliary heating. One arrangement of a cell element is shown in Fig. 13.18.

Figure 13.19 shows two performance curves. The platinum electrocatalyst is supported on C in one. The marked advantage of the more conducting TiC as support (less IR drop through the support) is shown.

The phosphoric acid cell has been under research for a longer time than that of any other kind of fuel cell. Alloys of Pt with Cr, V, and Ti and other non-noble metals are better than Pt (Appleby, 1986). The particle size of the catalyst has been reduced to that of tens of atoms (Stonehart, 1993).[10] Much attention has been given to the search for non-noble (hence cheaper) catalysts that are stable in hot acids. The best are the porphyrins, the formulas for which are shown in Fig. 13.20. They are applied to a base of graphite. These electrocatalysts are more effective in alkaline fuel cells than in those with acid electrolytes. Curiously, these substances are more stable and give better catalysis after pyrolysis in He at 800 °C, a process that would decompose the organic part of the structure. Perhaps the only active part of the porphyrin catalyst is the central

---

[10]In large particles, most of the expensive Pt atoms within the crystallites are not utilized for the electrocatalytic reaction.
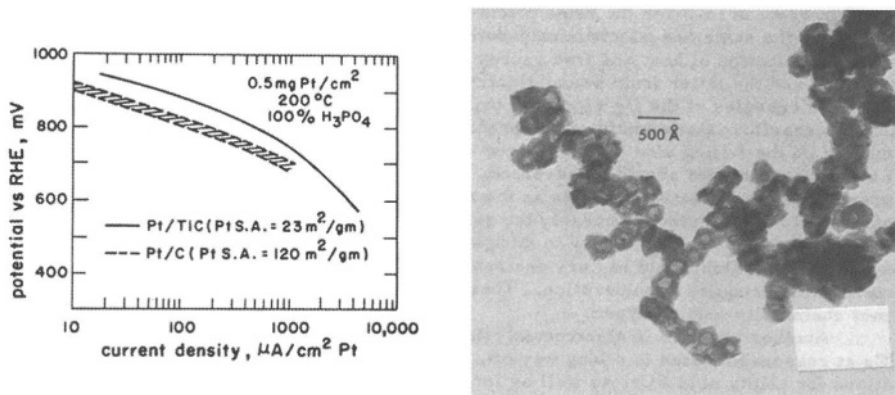
Fig. 13.19. Comparison of the activity of Pt for oxygen reduction on high-surface-area titanium carbide and on Pt on Vulcan XC-72R carbon black (see the TEM photographs of the TiC support). (Reprinted from *Assessment of Research Needs for Advanced Fuel Cells*, S. S. Penner, ed., Pergamon Press, 1986, pp. 14, 87.)

metal atom (see Fig. 13.20), and the organic structures serve to keep the individual atoms apart, preventing coalescence and growth to larger particles with the loss of an effective catalyst *inside* larger spheres (which does tend to occur otherwise). Pyrolysis, destroying the organic structure, bonds the atom to the underlying structure; that is, it also prevents coalescence and allows the continued activity of each atom in the catalyst.

The phosphoric acid fuel cell is being increasingly used to provide light and heat in large buildings. Use of the heat produced in the cell ("co-generation") is economically attractive. It increases the total use of the energy of the overall reaction from 40% to over 80%.

13.6.5. High-Temperature Fuel Cells

Compromise is the name of the game played by designers of high-temperature fuel cells. By raising the temperature to above 650 °C, there is a major advantage in reducing activation overpotential losses toward zero because the i_0's of the electrode reactions are greatly increased (cf. the Tafel equation). However, the high temperatures cause corrosion and loss of active materials and hence shorten the lifetimes of fuel cells. Molten carbonate fuel cells (600–700 °C) have been researched (Broers, 1960) for an extent of time second only to that for alkaline cells. Figure 13.21 shows a version of such cells, including reactions that would take place if the cell were to run basically on coal. Some carbonate fuel cells have been kept running continuously for more than 5 years. The efficiency of their production of electricity is greater than 55%. Industrial co-generation of electricity and heat is an ideal market, even though the current incentives in the United States, Japan, and Europe are geared toward electric utility power generators.
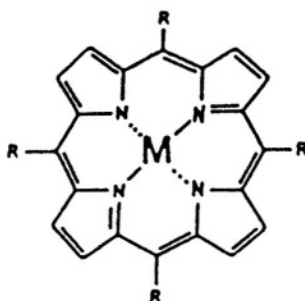
Fig. 13.20. Structures of porphyrins and iron porphyrins.

Certain oxides (particularly ZrO_2 stabilized by Y_2O_3) conduct ionically (O^{--} is the mobile ion) at temperatures of 1000 °C and above. Thus, solid ZrO_2 (Y_2O_3) is a "solid electrolyte" and can be used in a fuel cell. The cathode material is porous strontium-doped $LaMnO_3$ and the anode is Ni on ZrO_2. The advantage of the solid electrolyte (compared, say, with the liquid carbonate) is the reduction in corrosion resulting from the very low diffusion coefficients of ions (except O^{--}) in the solid

Coal Gasification:
$$C + H_2O \xrightarrow{(O_2)} CO + H_2$$

Anode:
$$H_2 + CO_3^= \longrightarrow CO_2 + H_2O + 2e^-$$
$$CO + CO_3^= \longrightarrow 2CO_2 + 2e^-$$

Overall Anode:
$$H_2 + CO + 2CO_3^= \longrightarrow 3CO_2 + H_2O + 4e^-$$

Cathode:
$$4e^- + 2CO_2 + O_2 \longrightarrow 2CO_3^=$$

Fuel Cell:
$$CO + H_2 + O_2 \longrightarrow CO_2 + H_2O$$

Overall Powerplant:
$$C + O_2 \longrightarrow CO_2$$
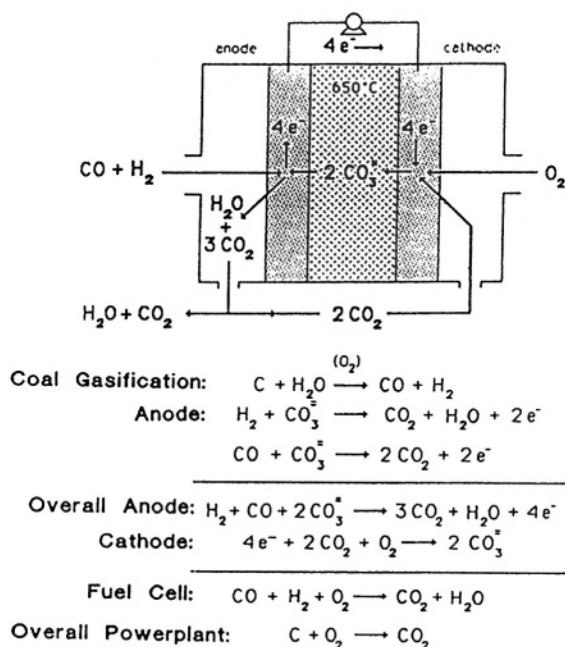
Fig. 13.21. Simplified schematic of a coal-fired molten carbonate fuel cell. (Reprinted from A. J. Appleby and F. R. Foulkes, *Fuel Cell Handbook*, p. 540, Fig. 17-1, copyright © 1993, John Wiley & Sons. Reprinted with permission of John Wiley & Sons, Inc.)

state. Because of the high temperature, re-forming to H_2 from CH_4 can occur within the cell without the need for a separate re-former.

   This solid oxide fuel cell may become the preferred version of the two high-temperature cells (Ackerman, 1983). More advanced designs involve lengthy passages in which fuel and air (O_2) follow parallel paths, each contacting an electrode, the two electrodes being separated by a long layer of ionically conducting ZrO_2 (the "solution") (Fig. 13.22). The ideal goal of such a "monolithic"[11] arrangement is to

---

[11]The use of the word "monolithic" (i.e., "a large organization that constitutes a uniform single unit") to describe certain fuel cell designs dates from 1983. However, Bockris and Srinivasan published a design in 1966 that demonstrated the basic idea of a pore active to an equal degree at all parts; they called this "spaghetti fuel cells." The necessary anodic and cathodic reactions were to occur on catalysts contained on the outside and inside of each tube, which were to be permeable to protons. They showed for the first time that many such tubes, put together in a single unit, would give very large power per unit volume because the whole of each tube would be active (compared with the porous electrode of conventional cells, where only the three-phase boundaries are active; Section 13.5.1).
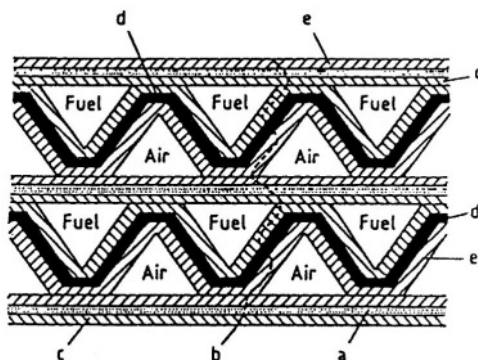
Fig. 13.22. The monolithic SOFC concept of Argonne National Laboratory. Anode: nickel-yttria-stabilized zirconia. Cathode: strontium-doped lanthanum manganite. Interconnect: doped lanthanum chromite. a, Interconnection; b, electron–ion path; c, anode; d, electrolyte; e, cathode. (Reprinted from K. Kordesch, *Ulmann's Encyclopedia of Industrial Chemistry*, Vol. A12, p. 82. Copyright © 1987 VCH-Wiley. Reproduced by permission of John Wiley & Sons.)

build very large cells (for there is no reduction of the activity "down the pore" to contend with). Such cells (called "monolithic"), well engineered, could accomplish much of the electricity production in technological societies with conversion efficiency at least 33% greater than that obtained at present thermal power stations.

13.6.6. Solid Polymer Electrolyte Fuel Cell

Sir William Grove in 1839 and Tom Bacon in 1932 each realized independently that water electrolyzers could be reversed in function to provide electricity when hydrogen and oxygen were introduced separately, each gas to its own electrode, with the electrodes separated by an aqueous solution. Such a beginning to electrochemical energy conversion implies that there will be an aqueous solution between the electrodes in fuel cells, and that is what exists now, e.g., in the alkaline fuel cell. When one thinks a bit further (what would happen if the cell were exposed to diminished ambient pressure, or to temperatures above 100 °C?), the feeling creeps upon one that it might be a good idea to get rid of the aqueous solution, but still work with H_2 and O_2.

How this might be achieved was first cleverly described by Grubb in 1957. In his idea, the H_2 fuel ionizes on a metal catalyst (Pt because the environment is acidic) and
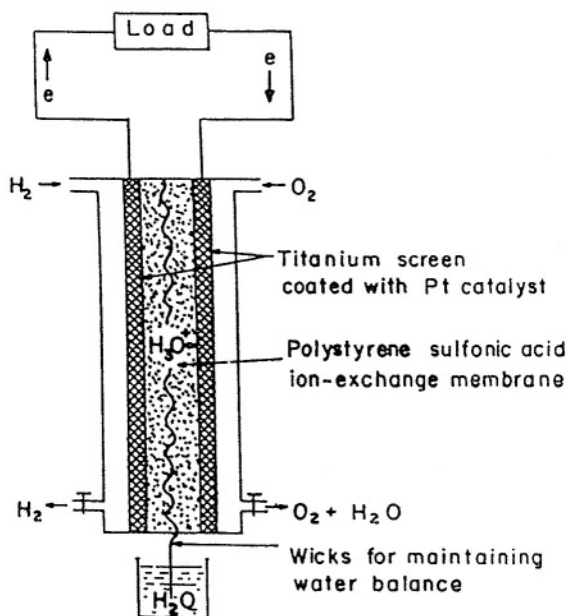
Fig. 13.23. Schematic representation of a single Gemini hydrogen-oxygen fuel cell.

the protons produced by the fundamental electricity-producing reaction $H_2 \rightarrow 2H^+ + 2e_0^-$ go through a thin (0.1-mm thick) membrane which, because it allows proton mobility (O^{2-} mobility in Section 13.6.5) is effectively "the solution" of the fuel cell. On the other side of the membrane is another electrode in contact with O_2 and here occurs the reaction $O_2 + 4H^+ + 4e \rightarrow 2H_2O$, the protons that the reaction demands being those coming through the membrane from the ionization of H_2 on the other side. Thus, no bulk liquid water in the usual sense[12] is needed for the cell to function (the membrane containing mobile protons), although water is produced as the end product at the cathode.

---

[12]However, water is vitally required for the solid polymer electrolyte fuel cell to operate. The membrane is a perfluorosulfonic acid polymer (e.g., Nafion) and for it to operate well as the electrolyte in a fuel cell, it has to have about 50% water. The membrane has a structure consisting of clusters and channels. The clusters have sulfonic groups, which face inside. The channels are composed of a completely fluorinated carbon chain. There are three types of water: the primary and secondary hydration sheaths for the H^+ and $C\text{-}SO_3^-$ ions, and water in the clusters and channels. The proton conductor acts in the membrane by a Grotthus mechanism, just as in an aqueous acid solution. The advantages of using Nafion-type membranes are: (1) Fluorinated acids are superacidic and have high conductivities. (2) Very thin membranes ($50\,\mu m$ or less) can be used for the electrolyte layer, which is most important in reducing ohmic losses. (3) The polymer electrolyte also serves as an excellent gasketing material for the cells.

Grubb's ideas can be shown schematically as follows (Fig. 13.23): The realization of this concept in practice depends on a suitable membrane, and this was found in Nafion $(CH_2)_{18-28}$ SO_3 H, which can be represented as

$$-(CF_2\ CF_2)_n - CFO\ (CF_2 - CFO)_m - CF_2 - CF_2 - SO_3 - Na^+$$
$$\quad\quad\quad\quad | \quad\quad\quad\quad |$$
$$\quad\quad\quad CF_2 \quad\quad\quad CF_2$$
$$\quad\quad\quad\quad | \quad\quad\quad\quad |$$
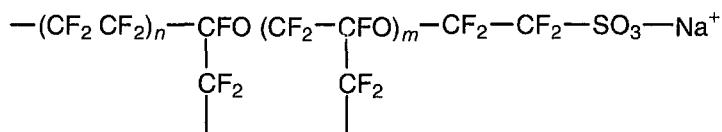
The description given here is a basic outline of the principles of the solid polymer electrolyte fuel cell used in the first Gemini space flights with nonfluorinated membranes (Fig. 13.23). Because the cell is slated for development as part of the electrochemical engine in cars, stages in its modern development are described in another section.

13.7. ELECTROCHEMICAL ENGINES FOR VEHICULAR TRANSPORTATION

13.7.1. The Electrochemical Engine

In the past, electric cars running on batteries have had the disadvantage of the weight of lead-acid batteries, the most developed and available battery. Considerable progress in the science and technology of batteries (Section 13.14) has been made, but the fact remains that the new (lighter) batteries also take several hours to recharge, and all except cells based on aluminum (primary batteries needing "mechanical" charging) provide ranges less than those of internal combustion-driven cars.

The term "electrochemical engine" refers to the fuel cell-electric motor combination and was introduced by Douglas Henderson at the GM Allison Division in 1967. It has been understood for several decades that the use of a fuel cell instead of a battery in electric cars would overcome the recharging and range problem associated with batteries (for there is no recharging and the range simply depends on how much fuel the car carries). However, research on the direct use of methanol or other hydrocarbon fuels for fuel cells had not reached a satisfactory development stage even in the 1990s, and economic photovoltaics (to provide the electricity to split water and produce hydrogen) seemed more than a decade distant.[13] For this reason, Billings (1991) suggested on-board re-forming of natural gas to hydrogen (the fuel cell fuel) as an interim measure until a massive supply of H_2 from dissociation of water could be engineered at an economic level. This suggestion, published in a textbook in 1993,

---

[13]If sufficiently cheap hydrogen from the solar-driven photoproduction of hydrogen were available, the hydrogen could be stored on board in lightweight cylinders or as a liquid (no re-former!).

was first put into practice by the Daimler-Benz Company of Germany (makers of the first passenger automobiles in the nineteenth century and at present manufacturers of the prestigious Mercedes car). In 1997 they exhibited three vehicles (a sedan, van, and truck) that were electrically driven and powered by a proton-exchange membrane fuel cell manufactured in Canada by Ballard Power Systems, Inc. This cell ran on hydrogen produced on board the vehicle by a re-former taking in methanol and using it to make H_2 and CO_2. The advantage of methanol over natural gas is that it is a liquid at room temperatures and can be stored in tanks, thus using the infrastructure of the present gasoline-based system.

Thus, the Daimler-Benz system has the advantage of taking on board a liquid fuel are retaining the present infrastructure. On the other hand, methanol (though easily manufactured from CO and H_2 produced by steam re-forming of hydrocarbons) is not yet available in the quantities needed, so that the fuel actually re-formed to hydrogen is likely to be gasoline, although the re-forming reaction here is less complete than that with methanol and needs a higher temperature.

The systems to be used in electric cars powered by fuel cells are shown schematically in Figs. 13.24a and 13.24b. The on-board re-forming of carbonaceous fuels to form H_2 has several marked advantages.

1. It diminishes the political opposition of the oil imperium. Oil and coal-derived products will still be sold as transportation fuels until the cost of photovoltaic hydrogen from water becomes competitive.
2. The economics should be attractive to the customer because of the increased efficiency of the electrochemical engine over the internal combustion motor.
3. The main advantage is the elimination of many of the pollutants (CO, unsaturated hydrocarbons and various nitrogen by-products, referred to for convenience as NO_x) are associated with the operation of the internal combustion engine, which is at present the principal source of smog in cities; the elimination of carcinogenic products; and the avoidance of SO_2 pollution (acid rain), which would have increased had batteries been chosen as the main power sources for electric cars. Thus, the widespread introduction of batteries into automotive transportation would have required a massive increase in the combustion of sulfur-containing coal to make the extra electricity for charging the batteries.

13.7.2. The Re-former

The re-forming of methanol to H_2 and CO_2 can be carried out by the following overall reaction:

$$CH_3OH + H_2O \rightarrow 3H_2 + CO_2$$
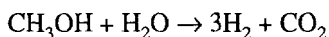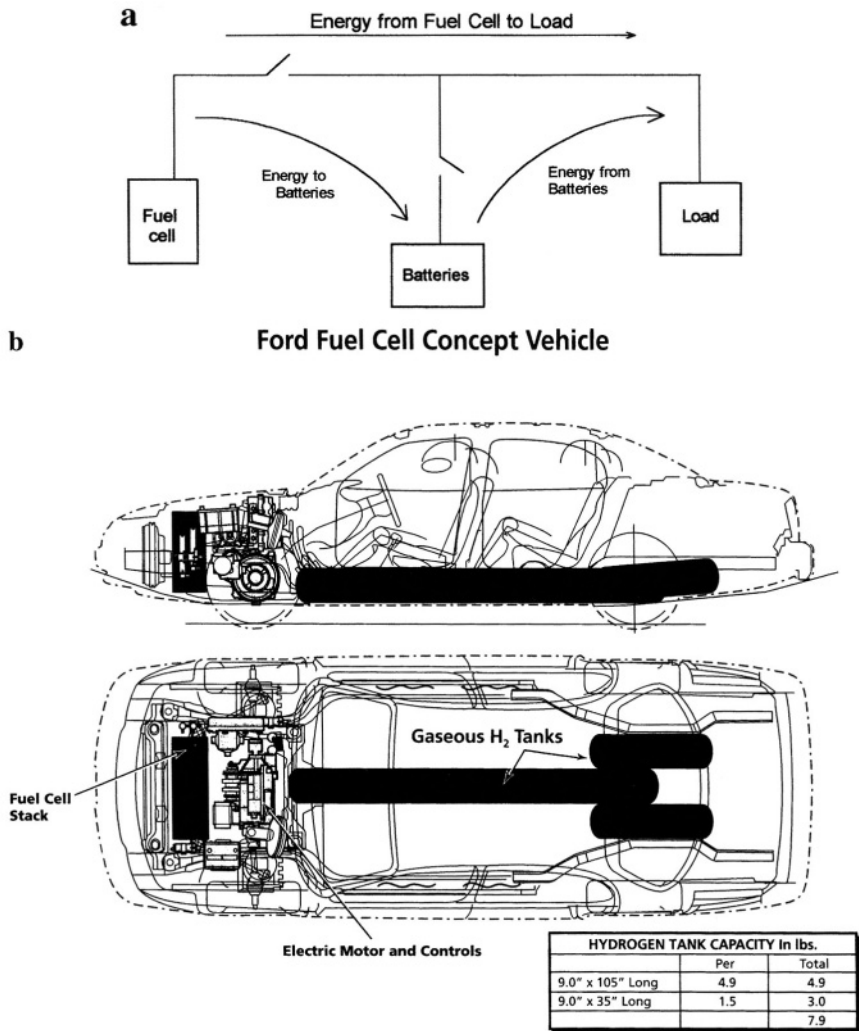
Fig. 13.24. (a) A simplified block diagram depicting the ability to selectively connect the fuel cell and/or battery pack to the load. (b) A Ford design for an H_2-fueled car. (Reprinted with permission from *Fuel Cells for Transportation*, U.S. Dept. of Energy, 1995, p. 14.)

However, if one writes the reaction in this way, a point is missed—the presence in the incompletely re-formed gas of the powerful electrode poison, CO. Thus, the re-forming reaction proceeds via the mechanism:

$$CH_3OH \xrightarrow{catalyst} H_2 + CO$$

$$CO + H_2 \rightarrow CO_2 + H_2$$
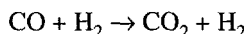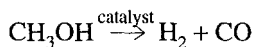
(the so-called "shift" reaction).

Unlike hydrocarbons and ethanol, which require temperatures of 800 °C, the re-forming reaction for methanol occurs on catalysts at a moderate temperature of 250 °C. The catalyst generally described for this reaction is CuO–ZnO together with a number of doping agents. The process is carried out by taking vaporized methanol and water and reacting them on the catalyst surface to form an $H_2 + CO_2$ stream for the fuel cell. It is essential that this stream be cleaned up before reaching the cell, for CO poisons the anode, and unreacted methanol can dissolve in the solid polymer of the fuel cell that is at the forefront of development as the power source for vehicular transportation.

If the CO is left untreated, it would reach the anode at a concentration of 0.5–4% within the mixture from the re-former. The preferred approach is to oxidize the CO further to CO_2. This oxidation must be made preferential to CO and avoid any oxidation of H_2, the fuel. The overall setup is shown in Fig. 13.25.

The Shell Company projects (1999) the advantages of re-forming *gasoline* with their patented "catalytic partial oxidation" process, which produces hydrogen from gasoline. The temperature needed (1000 °C) is still far above that for the re-forming of methanol. Whether this re-forming will compete with that of methanol at 250 °C has to be established. What of the completeness of the re-forming reaction? However, it has the huge advantage of re-forming the currently available gasoline.

The potential–current density relation for a PEM cell is shown in Fig. 13.26. One sees two results of interest. First, the figure shows the general trend for all fuel cells. As the current density increases, there is first a decrease in the available potential due to the normal overpotential associated with all electrode reactions, then a long quasi-linear section that represents increasing IR losses, and finally an approach to a limiting current at high current densities (see the simpler and more pronounced behaviors in Figs. 13.5–13.11 for flat electrodes). Thus, acceleration in a vehicular application is associated with a decrease in the efficiency of energy conversion (a buildup of overpotential with an increase in current density). The other point to note is the effect of the components of air on the performance (Fig. 13.26) of the anode in the fuel cell. There is an earlier departure from linearity in the cell potential vs. current density plot.
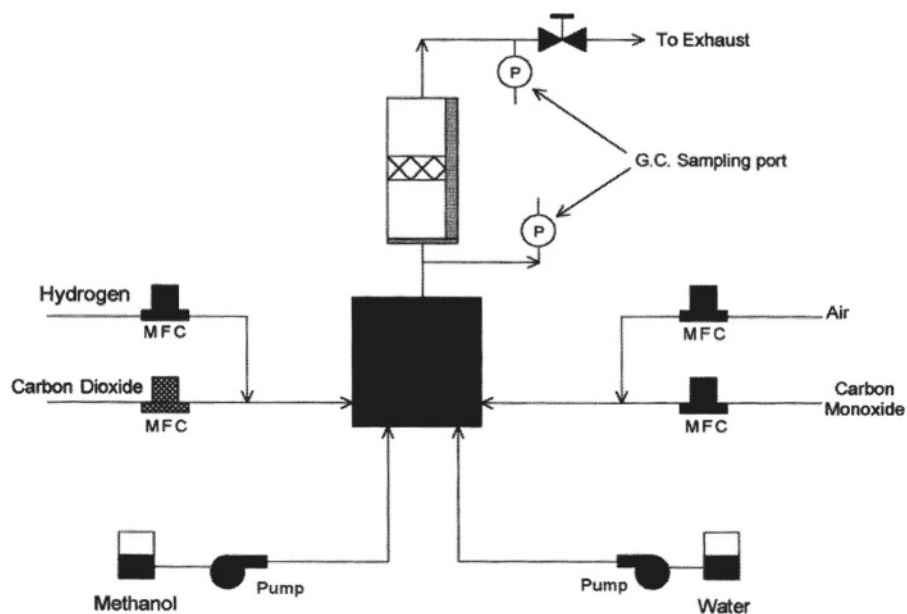
Fig. 13.25. Preferential oxidation and re-former–shifter test bench experiment. (Reprinted with permission from "Research and Development of Proton-Exchange-Membrane (PEM) Fuel Cell System for Transportation Applications, Phase I, Final Report," prepared for the U.S. Dept. of Energy by General Motors, 1996, Fig. 2.2.3.1.)

13.7.3. Development of the Proton-Exchange Membrane Fuel Cell for Use in Automotive Transportation

13.7.3.1. *General.* There is no doubt that a comparison (updated in the 1990s) of the potential vs. current density plot for the various fuel cells (see Fig. 13.27) shows that the proton exchange membrane fuel cell with a perfluoropolymer sulfuric acid has superior performance (i.e., higher cell potential and hence efficiency) compared with the other types of fuel cells. Because this cell has been chosen for development by the majority of the automotive manufacturers, special attention is given here to its development.

13.7.3.2. *Fundamental Research that Underlay Development of this Cell.* Three U.S. universities were involved in the work that culminated in manufacture of the proton-exchange membrane by Ballard Power Systems. First, Case-Western Reserve University must be recognized because of the sustained investigations there (Yeager et al., 1961–1983) on the mechanism and catalysis of the reduction of O_2, the reaction that causes most of the energy losses in the fuel cell. *The Electrochemistry of*
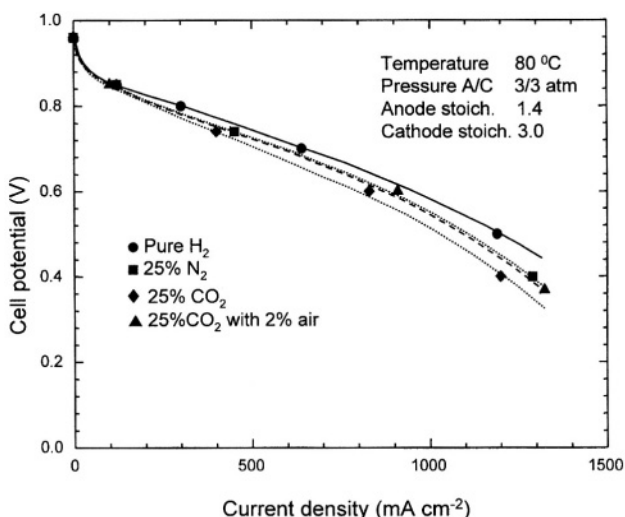
Fig. 13.26. The effect of reformate and reformate with 2% air anode feed on the voltage-current characteristics of a Dow membrane in the reference cell. (Reprinted with permission from "Research and Development of Proton-Exchange-Membrane (PEM) Fuel Cell System for Transportation Applications, Phase I. Final Report," prepared for the U.S. Dept. of Energy by General Motors, 1996, Fig. 3.4.5.5.)

Fuel Cells (Bockris and Srinivasan, 1969) has been the source book used in many investigations of fuel cell reaction mechanisms and it was written largely at the University of Pennsylvania and partly at the Down State Medical Center of the State University of New York. The University of Pennsylvania was the site of the fundamental work on the O_2 reduction mechanism which led to the conclusion that the rate-determining step (in acid solution on Pt) is $O_2 + H^+ + e \rightarrow O_{ads} + OH_{ads}$ (under Temkin conditions). This work (Damjanovic and Brusic, 1967) had much influence on the field through its translation from the university laboratory to fuel cell companies carried out by A. John Appleby (Texas A&M University), who became the most influential academic in the fuel cell area, with a world-wide influence on the field.

Damjanovic and Genshaw's 1970 paper experimentally establishing a *linear* relation between the coverage of the electrode by adsorbed O and the electrode potential had the effect of establishing the reality of Temkin conditions (Sec. 7.7.3). Another influential paper of these early days was the summary of design indications for fuel cells (Srinivasan and Cahan, 1968) applying to a great degree the model calculations performed by Srinivasan and Hurwitz (1967) and those by Bockris and
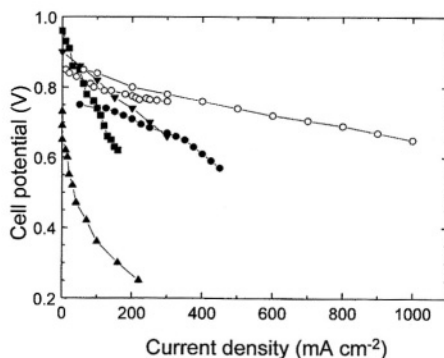
Fig. 13.27. Potential vs. current density plots for state-of-the-art fuel cells. ○, proton exchange membrane fuel cell; •, solid oxide fuel cell; ▼, pressurized phosphonic acid fuel cell (PAFC); ▲, direct methanol fuel cell, direct methanol PAFC; ■, alkaline fuel cell. (Reprinted from M. A. Parthasarathy, S. Srinivasan, and A. J. Appleby, "Electrode Kinetics of Oxygen Reduction at Carbon-Supported and Unsupported Platinum Microcrystallite/Nafion Interfaces," *J. Electroanalytical Chem.* **339:** 101–121, copyright 1992, p. 103, Fig. 1, with permission from Elsevier Science.)

Cahan (1969), all work carried out in the Chemistry Department of the University of Pennsylvania.[14]

At Texas A&M University, contributions by Srinivasan and co-workers can be connected in a direct line to the successful developments at Ballard. One of the more basic conclusions arising from Cahan's calculations of the distribution of current in the pores of the fuel cell concerned the limited number of pores in a porous electrode of a fuel cell that is actually used. Hence, the practice of distributing the Pt catalyst uniformly throughout a porous electrode meant that much of the Pt was ineffective. This conclusion was tested at Texas A&M's Center for Electrochemical Systems

---

[14] Copies of Boris Cahan's thesis were circulated among many working in the fundamentals of the field in the 1970s. The thesis gave its readers access to the details of the elegant experimental studies it reported on the interplay between meniscus shapes and the thickness of the boundary layer, i_0, electrolyte conductivity, and the local heat developed in porous electrodes used in fuel cells (and which tends to dry out thin menisci).

(Srinivasan, 1988). The resulting report showed that restricting the platinum to near the front of the electrode gave an enhanced utilization factor (50%), thus reducing the amount of Pt needed to 0.1 to 0.2 mg cm^{-2}. Very thin layers (100 μ) were achieved in PEM fuel cells (Srinivasan, 1990), and high current densities (> 2 A cm^{-2}) were achieved by 1993 in work at the same institution (see Fig. 13.28).

Thin catalyst layers made by using platinum on carbon electrocatalysts with a high Pt/C weight ratio (the Pt concentrated near the front surface) were found to be a principal factor in reducing the overpotential and allowing apparent current densities of more than 2 A cm^{-2} under special conditions (pressurized oxygen) to be developed with a loss of only 0.3 V from the reversible potential. Since the efficiency of the cell is proportional to ($E_{rev} - \Sigma\eta$), this is a significant achievement having a direct effect on the efficiency of energy conversion and hence the cost of transportation. The particle size distribution in these thin-layer PEM cells, first built at Texas A&M, is shown in Fig. 13.29. If the power density of such cells is calculated in terms of kilowatts per gram of Pt, there is an increase of up to 0.1 mg cm^{-2}, as shown in Fig. 13.30, which may be the optimum quality to be used in a fuel cell electrode.
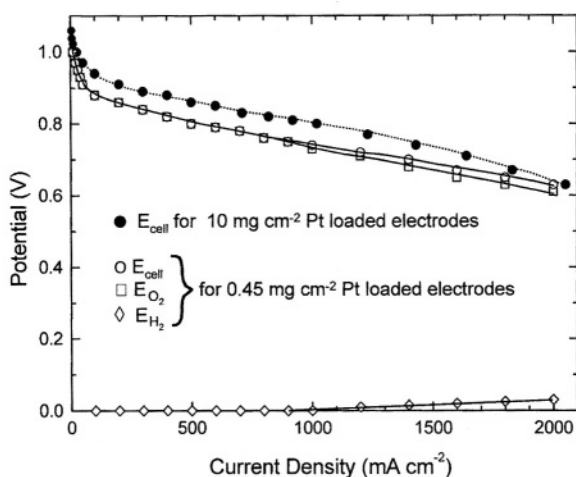
Fig. 13.28. TEM micrograph of 20 wt.% Pt-C electro-catalyst powder (Prototech). Vulcan XC-72 carbon support electrodes are now 0.2 mg cathode, 0.05 (for H$_2$) anode. (Reprinted from M. A. Parthasarathy, S. Srinivasan, and A. J. Appleby, "Electrode Kinetics of Oxygen Reduction at Carbon-Supported and Unsupported Platinum Microcrystallite/Nafion Interfaces," *J. Electroanalytical Chem.* **339**: 101–121, copyright 1992, p. 105, Fig. 2, with permission from Elsevier Science.)
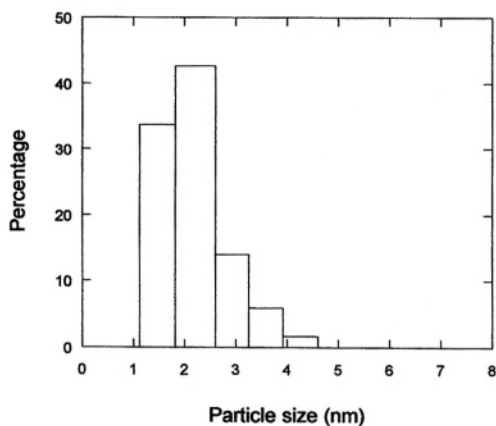
Fig. 13.29. Distribution of particle sizes in 10 wt. % Pt-C electrocatalyst. Particle sizes are average diameters. (Reprinted from E. A. Ticianelli, M. N. Beery, and S. Srinivasan, "Dependence of Performance of Solid Polymer Electrolyte Fuel Cells with Low Platinum Loading on Morphologic Characteristics of the Electrodes," *J. Appl. Electrochem.* **21:** 601 copyright 1991, Fig. 9.

The development of a satisfactory power source for electric cars toward the century's end has been a huge effort, involving laboratory work in 12 states. It is funded by the U.S. government through the Department of Energy, together with substantial input by the three automotive companies, under a program called "Partnership for a New Generation of Vehicles." In Europe, Daimler-Benz is doing intense (and, in terms of time, leading) work in this area (although, as indicated, the fuel cell is manufactured in Canada). In Japan, the Toyota company is second only to Daimler-Benz in planned time to the commercialization of a fuel cell-driven electric car.

The key achievements at Ballard (Wilkinson 1998)[15] are low Pt loading (1 mg cm^{-2}) in 50-kW fuel cell power plants, and that of a power-to-weight ratio of 1 kW/kg. The development of the solid polymer electrolyte membrane was by no means in a final state in the late 1990s. The quality of the membrane controls the highest current density at which the cell is viable. There are open areas, too, with regard to the composition (as apart from the loading and particle size) of the catalyst; a PtRu alloy

---

[15]It is of interest to note that Dr. D. P. Wilkinson, who leads the electrochemical work at Ballard, obtained his graduate training with Prof. B. E. Conway, a member of the same electrochemistry group from which Rex Watson originated. Watson was the sole electrochemical researcher on Bacon's team in 1952.

Pressure 60/60 psig - H_2/O_2
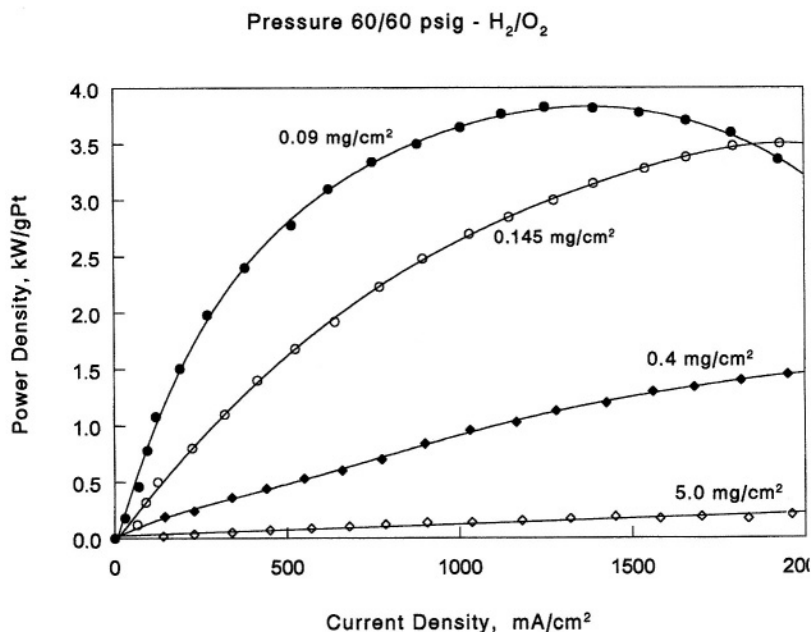
Fig. 13.30. Power density in kilowatts per gram Pt as a function of current density for several degrees of application of the catalyst in milligrams per square centimeter. (Reprinted from S. Srinivasan 28th Intersociety Energy Conversion Engineering Conference, American Institute of Aeronautics and Astronautics, 1993).

offers a substantial improvement over Pt, but the possible advantages of multisite catalysts (Minevski, 1992) seem not to have been explored yet.

The efficiency of a practical 50-kW fuel cell (International Fuel Cells, 1997) as a function of power density is shown in Fig. 13.31 and can be compared with the schematics of flat plate fuel cells (Fig. 13.9, where the efficiency is plotted against current density).

13.7.4. The Electric Car Schematic

The arrangement of the electric car as seen by the U.S. Department of Energy in 1996 is shown in Fig. 13.32.

13.7.5. A Chord of Continuity

Some reference has been made to the university work (by Yeager, Damjanovic, Cahan, Srinivasan, and Appleby) that has formed the basis for the jump in fuel cell
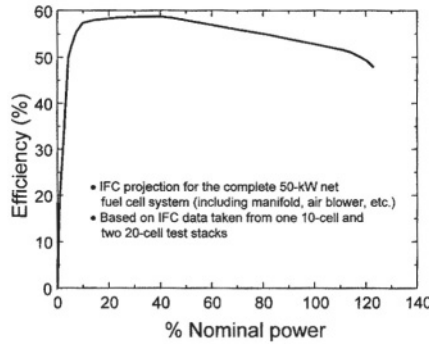
Fig. 13.31. Fuel cell efficiency vs. power (International Fuel Cells 50-kW$_e$ PEM fuel cell system). (Reprinted from *Fuel Cells for Transportation*, U.S. Dept. of Energy, 1996.)

development in the 1990s. However, there are closer and more continuous links than those stated, and these bind Bacon to Watson (from Imperial College, London University) and the work of the latter in 1952 on the mechanism of the hydrogen electrode reactions in alkaline solutions. B. E. Conway (later the doctoral advisor for Wilkinson of Ballard), had preceded Walton in the same group at London University and in 1956 he authored a seminal paper relating the mechanisms of hydrogen reactions on electrodes to the electronic structure of the metal.
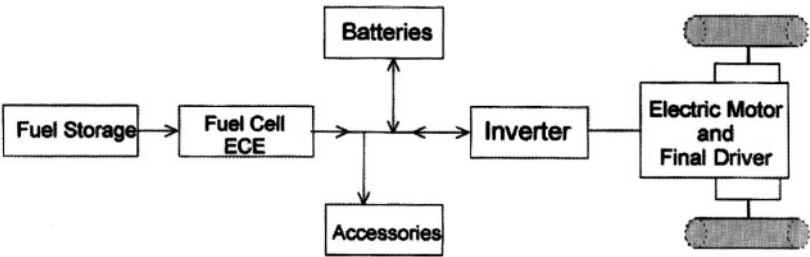
Fig. 13.32. Series fuel cell–battery hybrid power train model. The two-way energy flow between the drive motor and electric bus indicates the potential for regenerative braking. (Reprinted with permission from "Research and Development of Proton-Exchange-Membrane (PEM) Fuel Cell System for Transportation Applications, Phase I. Final Report," prepared for the U.S. Dept. of Energy by General Motors, 1996, Fig. 3.5.2.1.)

The University of Pennsylvania was the site of Damjanovic's work with Genshaw and Brusic (1966–67); in the same year from Pennsylvania came a design paper on fuel cells and electrode kinetics, co-authored by Srinivasan[16] (including calculations on a "spaghetti fuel cell"). Srinivasan began then his most active period at Pennsylvania, coauthoring with Wroblowa the first review of electrocatalysis (1967); publishing a theory of porous electrodes; and coauthoring the well-known McGraw-Hill publication on the electrochemistry of fuel cells.

Not to be forgotten is the work that underlay even this basic work on the present success in fuel cells: the textbook (1970) of which this volume is the second edition; and a founding paper on the concept of a hydrogen economy (with Appleby, later the director of the Center for Electrochemical Systems and Hydrogen Research at Texas A&M University). This center—which gave rise to much progress in fuel cell research directly preceding the work at Ballard—originated from the National Science Foundation's Hydrogen Research Center (1982–87) at Texas A&M University, where the central aim was the photosplitting of water to yield hydrogen as the fuel for fuel cells (no injecting of CO_2 into the atmosphere).

13.8. HYBRIDS INVOLVING FUEL CELLS, BATTERIES, ETC.

A number of hybrid schemes for electrochemical energy conversion have been devised. These include the use of a fuel cell to compress air, which would drive air turbines to provide startup and acceleration.[17]

---

[16]Supramaniam Srinivasan began his career as a customs officer in Sri Lanka, but came in 1960 to the University of Pennsylvania where he wrote a Ph.D. thesis on a new way to determine the path of the hydrogen dissolution reactions on electrodes. Several of his further contributions to the fundamental basis of fuel cells are described in our text.

The late 1960s saw "Srini" in his bioelectrochemical period at the Down State Medical Center, State University of New York, during which he determined (with P. Sawyer) the electrochemical mechanisms of arteriosclerosis (Sec. 14.9.3). A lengthy and productive phase (with a large team of collaborators) at the Brookhaven National Laboratory brought him to the later 1970s and was followed by positions in the Hydrogen Economy Institute in Canada and in the electric group at Los Alamos National Laboratory.

Srinivasan had contributed much to knowledge of electrochemical energy conversion in these several appointments when he found the most opportunity-filled position in the Center for Electrochemical Studies and Hydrogen Research at Texas A&M University. The director of the center, A. John Appleby, found that his position turned him into a world traveller to such an extent that Srinivasan became, in effect, the day-to-day director of the center.

Direct, intelligent, striving—these are words that come to mind in considering this most outstanding scientist. He met the stresses of a not always obedient body by a steely determination to rise again and again to give more to his field.

While Daimler-Benz and Ballard deserve public thanks for their plans to reduce planetary warming, those who know the history of this effort wonder how much longer it would have taken without the research direction provided by Dr. Srinivasan.

[17]In the Flinders University (Australia) electric car of 1976, a centrally located electric motor drove hydraulic fluid to turbines on each wheel.

13.9. DIRECT MEOH FUEL CELLS

Most fuel cell research activities are based on H_2, as the fuel delivered to the anode. The reason for this is that hydrogen dissolves electrochemically to form protons or water with an i_0 greater than that of any competing fuel. Although (see Chapter 15) an economic supply of H_2 from water using photovoltaic electricity may be a likely eventual result of U.S. and Swiss research, the prospects for some time yet are that H_2 will continue to be produced from the re-forming of fossil fuels, probably from gasoline itself (although the re-forming reaction of methanol runs at a much lower temperature. However, the re-forming stage is expensive and involves a weighty addition to the engine. The question therefore is: Could MeOH be directly transformed to electricity in a fuel cell?

At first, the prognostications for this seem poor. In the 1960s a number of oil companies attempted to oxidize propane, diesel oil, and alcohols in fuel cells at <100 °C. The i_0 for the anodic oxidation of simple saturated hydrocarbons is about 10^{-9} A cm^{-2}, at 80 °C, which is no greater than the rate of the sluggish cathodic reduction of O_2. If the anodic oxidation reaction of such hydrocarbons could be coupled with the cathodic reduction of O_2, the total overpotential of the fuel cell at reasonable rates would be so large that the efficiency gains expected over that of internal combustion engines would be doubtful. Since 1990, however, much progress has been made in the electrochemical oxidation of MeOH (Chapter 7). A difficulty in this process is the buildup of organic radicals on the surface of the electrode. Much depends on the time of electrolysis (minutes? hours? months?). After some time, a nonreactive CO species forms and later more complex unreactive organics, which are often called by the expressive term "gunk." Several studies (Enyo, 1988; Hamnet, 1995) have shown the effectiveness of some alloys of Pt ($Pt\text{-}Wo_3$; $Pt\text{-}Ru$; and rare earth tungsten bronzes doped with Pt) in controlling the buildup of the unreactive form of CO (Hamnet and Troughton, 1992). The use of ternary alloy electrocatalysts (Vielstich, 1993) has proved to be the most effective; Pt-Ru-Cr and Pt-Ru-Ni are used. The electrocatalysts are dispersed on carbon and can be used at temperatures up to 75 °C. They give hope of an acceptable direct methanol-fueled fuel cell. A fuel cell burning methanol directly will have to have a steady-state performance in which the CO_{ads} radicals do not build up to impede the reaction. Repeated programmed anodic pulses to clean the surface of such impeding materials may be helpful. XPS studies (Hamnet 1987) carried out on Pt-formed crystallites deposited on graphite have indicated that the key to high activity is the small size of the crystallites; it seems that such particles have enhanced electrocatalytic properties, due probably to some aspects of their single crystallinity.[18] The subject of the direct MeOH fuel cell is clearly one needing high

---

[18] P. Stonehart has specialized in methods for the manufacture of particles of a few atoms each. The art concerns less the actual manufacture of such particles than getting them to stay on the surface without undergoing "Ostwald ripening," the thermodynamic tendency for small particles to aggregate and grow in size, thus diminishing the fraction of the Pt that can be used in catalytic reactions.
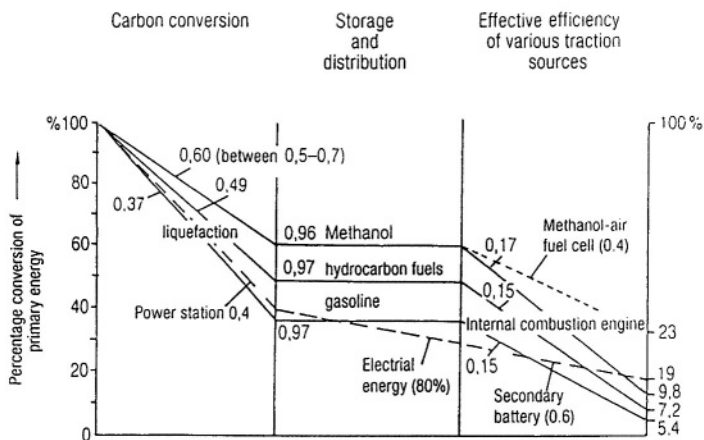
Fig. 13.33. Simplified energy-efficiency diagram for coal as the primary fuel for traction through liquefaction and electricity generation. The figures on each line are the approximate efficiencies for each process. The value of 0.8 for the storage and distribution of electrical energy contains the contribution from the efficiency of charging equipment. The value of 0.6 for the secondary battery (accumulator) combines the efficiency for storage (0.67) and electrical drive (0.9). (Reprinted from C. H. Hamann, A. Hamnett, and W. Vielstich, *Electrochemistry*, p. 382, Fig. 9-26, 1998, with permission from Wiley-VCH.)

intelligence in the electrocatalytic work. The relative efficiency of this cell compared with other power sources (Hamnet, 1998) is shown in Fig. 13.33. Srinivasan and Lamy have described the electrochemical kinetic aspects of methanol oxidation in a review in *Modern Aspects of Electrochemistry,* Vol. 34.

13.10.   GENERAL DEVELOPMENT OF A FUEL CELL-BASED TECHNOLOGY

13.10.1.   Fuel Cell Power Plants

Power plants using fuel cells can now take the place of the present polluting coal or oil-based (indirect) electricity-producing plants. However, in a further development, it would be possible to extract CO_2 from the atmosphere, and H_2 from solar-driven electrolysis, to produce methanol with zero net injection of CO_2 into the atmosphere. These plants would at first run on hydrogen from these fossil fuels, the attraction being the reduction of pollution and the increase in the conversion efficiency. To what extent the latter two commodities would be supplied from remote sites, or collected onsite at

the power plant, will be a matter of economics and depend heavily on the geographic location of the collector site and the average yearly solar intensity. If the CO_2 is extracted from the atmosphere (Stücki, 1996), the on-board re-forming of methanol (which rejects CO_2 for the atmosphere) need involve no net contributions to greenhouse warming. The actual fuel cell for large-scale power plants is likely to be a monolithic solid oxide cell.

13.10.2. Household Energy

One possibility for supplying household energy is to distribute electricity from central fuel cell-based power plants to houses in the surrounding area. However, it may become cheaper to store methanol in each plant and use it in the co-generation of heat and electricity.[19] Such a scheme would also make possible advantages in the distribution of lighting in households via pipes from a central light source powered by fuel cells. This type of situation may provide an application for phosphoric acid cells.

13.10.3. Vehicular Transportation

The use of the solid polymer electrolyte fuel cells for cars (running on hydrogen re-formed from gasoline) seems assured (Section 13.9); the principal advantage is an elimination of the smog-causing unsaturated hydrocarbons, and NO_x produced by internal combustion, with a 50% reduction in CO_2 and a consequent halving of the amount of fossil fuel that needs to be purchased (on the basis of a doubling in the efficiency of energy conversion).

Considerable advantages would accrue if it were possible to use pure hydrogen to power vehicles. The German BMW company proposes to do this and has demonstrated electric cars using liquid hydrogen as a fuel with a range of ~1000 km. In this scheme, the hydrogen would be manufactured at gas stations, at first from a delivered fossil fuel, liquefied, and taken on board (cryogenic storage). Without the need for an on-board re-former, the alkaline fuel cell (AFC) becomes most economic (as low as \$100/kW), although this may change with the parallel development of other kinds of fuel cells. If AFCs are to be used, air for the cathodes must have a CO_2 removal device because of carbonate formation.

13.10.4. Railways

The high-temperature solid oxide cell would be suitable for electric locomotives, with on-board re-forming of methanol or diesel oil within the cell.

---

[19]Overpotential causes waste heat to be developed when a fuel cell makes electricity at a significant rate. Why let this heat be dissipated into the atmosphere? Let us duct it to rooms in a building and raise the overall efficiency of conversion of the chemical energy in the fuel (i.e., heat and electricity) to 85%?

13.10.5. Seagoing Vessels

For cargo and passenger boats, conversion to fuel cells and electric drive will follow as a consequence of the more favorable economics of the lessened fuel consumption. Hence, the rate-determining step is the manufacture of sufficiently large fuel cells; once more the prospects look good for the monolithic solid oxide cell.

For naval vessels, an energy depot ship could use an on-board nuclear reactor to electrolyze brine (Sec. 15.3.5) and produce liquid hydrogen as a fuel for its accompanying fleet. An increase in the fraction of naval vessels that are submersible now seems likely. Thus, it would be advantageous to run on LH_2 under water for long periods with only water as the effluent. Conventional submarines operating on batteries can move under water for only a few hours before they have to surface to recharge. LH_2-fueled alkaline fuel cells could power non-nuclear submarines for times depending only on the size of the fuel tanks.

13.10.6. Aircraft

Airborne vehicles for passenger and cargo transport will be run on LH_2-fueled jet engines because of the 10 kW/kg power-to-weight ratio provided by these transducers. There may be possibilities for lighter-than-air vehicles with much area for photovoltaic cells on the large exposed surface of the vehicle. Here, on-board electrochemical cells could use water, electrolyze it to H_2, use this to power electrochemical engines, and stay aloft indefinitely (the craft would move at night on the H_2 stored from the day's electrolysis). The water electrolyzed to provide hydrogen for the fuel cells that would drive electric motors would be re-formed as a product of fuel cell action. Solar-powered stratospheric pilotless aircraft using fuel cells are being developed by AeroViron- ment, Inc., Monrovia, California, for upper atmosphere research. Lighter-than-air vehicles with 600 tons of carrying capacity are under development in Russia.

13.10.7. Industry

Power for industrial operations is supplied at present by electricity from coal, natural gas, and oil, working heat engines at overall efficiencies of < 35%. All can be replaced by fuel cells of various kinds operating for the next two to three decades on hydrogen from re-formed fossil fuels; and as soon as possible by hydrogen from photovoltaics and the decomposition of water.[20]

---

[20] There are advantages in the suggestion (Stücki, 1996) that CO_2 could be extracted from the atmosphere by absorption in KOH and electrolysis of the carbonate, thereafter using chemical catalysis to convert CO_2 to methanol. It would make us independent of fossil fuels! No net increase or decrease of atmospheric CO_2 would occur because the absorption of CO_2 from air would be made up by the evolution of CO_2 from the re-forming of the MeOH. However, the electrolysis of carbonate gives one CO_2 molecule and one H_2 molecule, whereas the catalytic formation of methanol from CO_2 and H_2 needs two more H_2 molecules per CO_2 molecule. Thus, forming methanol from air-based CO_2 would still need the formation of H from water using renewable energies, as well as the use of electricity from renewables for the carbonate decomposition.

13.10.8. Space

The use of fuel cells on the Gemini series (and all subsequent) space flights run by NASA is well known. However, there are no diurnal variations of solar light in space, so that photovoltaics can provide the power for most space stations (as in the Russian "Mir") and on longer space flights.

13.11. THE SECOND FUEL CELL PRINCIPLE

The normal meaning attached to the phrase "the fuel cell principle" describes the addition of two chemicals to a fuel cell to produce electricity. However, it is a fact of electrochemistry that the sum of the two electrode reactions in a fuel cell amounts to a chemical reaction: a chemical product is synthesized. Now, all fuel cells in operation at this time use the cathodic oxygen reduction reaction $O_2 + 4H^+ + 4e \rightarrow 2H_2O$; and the anodic dissolution reaction $H_2 \rightarrow 2H^+ + 2e.$ Doubling the latter reaction and adding to that of oxygen reduction gives a *chemical* overall reaction $O_2 + 2H_2 \rightarrow 2H_2O$; i.e., the synthesis of water takes place in the necessarily spontaneous working of the cell. The news that one can synthesize water is not one to stir great interest in Wall Street (though it is useful in space stations) and our familiarity with water and the ease with which it is available hides the second fuel cell principle. Thus, the first fuel cell principle concentrates on the production of energy, but neglects the substance also produced.

Suppose, however, one reversed the emphasis and made a second fuel cell principle concerning spontaneous electrosynthesis. Then one would concentrate, not on the energy produced (now the by-product), but on the substance. Obviously, one would have to choose a situation in which the spontaneously acting overall reaction in the fuel cell produced a worthwhile product. To take a very simple example, suppose one led Cl_2 gas, instead of O_2 to the cathode of a fuel cell and ethylene instead of hydrogen to the anode. At the cathode one would find:

$$Cl_2 + 2e \rightarrow 2Cl^-$$
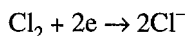
At the anode:

$$C_2H_2 + 2Cl^- \rightarrow C_2H_4Cl_2 + 2e$$
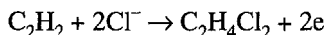
One has synthesized a new organic compound, a useful one, but at the same time there is a by-product, electricity. The electrogenerative synthesis of dichlorethylene has been known for several years. One might consider one of the most needed and best-known chemical syntheses, the ammonia synthesis, an exothermic reaction with a heat of reaction of 46.38 kJ.

$$3H_2 \rightarrow 6H^+ + 6e$$

$$N_2 + 6H^+ + 6e \rightarrow 2NH_3$$
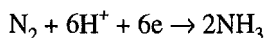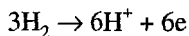
The cathodic reaction needs study as to the pressure and temperature range of the electrocatalyst.[21] Could it be accomplished at ambient pressure? (The expensive Haber chemical synthesis occurs at 10^2 to 10^3 atm and needs 500 °C for sufficient speed, together with a complex Fe catalyst.)

Now, in the first fuel cell principle ("the cell produces electricity from chemicals"), one neglected the substance produced. In the second fuel cell principle ("the cell produces chemicals electrochemically without the use of outside electricity"), one sees that the by-product is not the negligibly important water, but electricity. Compare electricity with heat, the by-product of our present indirect thermal system of producing electricity. The Carnot principle makes the efficiency of the heat engine quite low, say, 33%. So, 67% of the chemical energy used in operating a heat engine to create mechanical work may be given out as heat. But the properties of heat are very different from those of electricity, for the latter can be moved anywhere, for example, sent to a town 400 km away. Maybe the heat, if turned into steam, can be piped around some areas of a town. But no one suggests that heat has the same transmission properties as electricity. So, the second fuel cell principle is important because when it is used to make compounds, it not only produces them without the need to buy energy, but gives us an electricity supply as well. Figure 13.34 shows a practical example of this, the use of the fuel cell principle to make oxidation products of toluene.

How far could one go with this second fuel cell principle? In principle, any spontaneous chemical oxidation or reduction reaction can be simulated in a fuel cell, whereupon the cell becomes an electrochemical reactor. In the oxidation reactions, the coupling reaction is the cathodic reduction of oxygen. The oxygen can come from air at no cost. In the reduction reactions, one would tend to use H_2 at the anode if it were available cheaply enough, or maybe a cheap organic fuel, or even, with sufficient electrocatalysis and a rise in temperature, cheap biowastes on their electrochemical way to CO_2 (see Chapter 15). Now that the transport system is to run using the first fuel cell principle, it is worthwhile peering a bit into the future, and asking whether the second fuel cell principle has the power to change the way industry manufactures chemicals (Fig. 13.35). This route has the attractions of reducing polluting effluents and avoiding harmful, toxic by-products as well as being electrogenerative.

The concept of a nonpolluting (electrogenerative) chemical industry is a fine one, not yet realized, but perhaps only one generation in the future. The schematic in Fig. 13.35 illustrates the idea.

---

[21]During the electrochemically driven reduction of $NaNO_3$ at room temperature in an alkaline solution, the percentage of NH_3 in the gaseous products reached 60 when an Fe cathode was used (Kim, 1997).
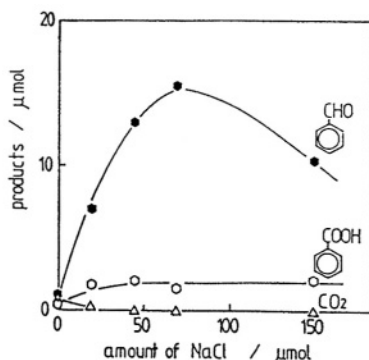
Fig. 13.34. Electrogenerative production of two organic compounds by a fuel cell approach. (Reprinted from K. Otsuka, K. Ishizuka, I. Yamanaka, and M. Hatano, "The Selective Oxidation of Toluene to Benzaldehyde Applying a Fuel Cell System in the Gas Phase," *J. Electrochem. Soc.* **138**: 3178, 1991, Fig. 3. Reproduced by permission of The Electrochemical Society, Inc.)

Now let us consider a 1-year period during which a fuel cell plant consumes 10,000 tons of methanol. This is 10^{10} g, and since the molecular weight of methanol is 32, the number of moles of methanol consumed in a year would be about 3×10^8. However, each mole of methanol in an anodic oxidation produces 6 F of electricity $(CH_3OH + H_2O \rightarrow 6H^+ + CO_2 + 6e)$.

Hence, the total number of coulombs produced in a year from 10,000 tons of methanol would be $3 \times 10^8 \times 6 \times 10^5$ C (since 1 F $\simeq 10^5$ C). Thus the oxidation of methanol in a fuel cell (10,000 tons per year) would produce 1.8×10^{14} C.

Let us suppose that the total current from all the fuel cells used in the electricity-generating plant is I A. The number of seconds in a year is 3.1×10^7 and (as It is coulombs since t is the time in seconds), $I = 3.1 \times 10^7 = 1.8 \times 10^{14}$ or $I = 6.6 \times 10^6$ A. Now, all this current would be converted to electrical energy in the fuel cells at (say) about 0.7 V, which is a reasonable potential in the oxidation of methanol in a fuel cell with a good electrocatalyst. Hence, we should produce 4×10^6 W or about 4000 kW.

Houses go through a cycle in their daily use of electricity, but the only value used here is the average power needed by a typical household leveled out throughout the year. If this is 4 kW, then about 1000 houses could be supplied with electricity from the imaged fuel cell plant.

The Second Fuel Cell Principle In Action
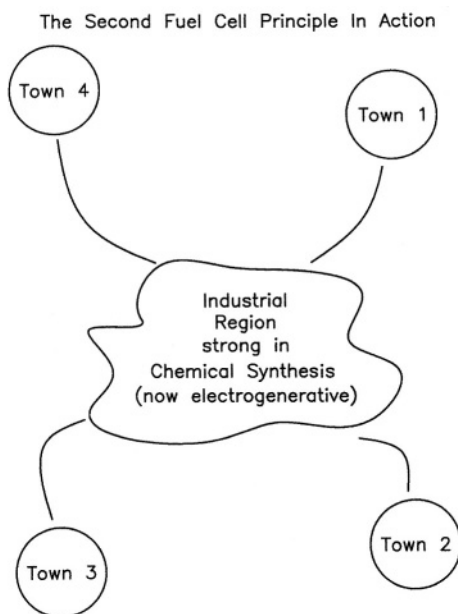
Fig. 13.35. A speculative schematic of a possible future application of the second fuel cell principle.

13.12. MIDWAY: THE NEED TO REDUCE MASSIVE CO_2 EMISSIONS FROM MAN-MADE SOURCES

The decision to power the transportation system with fuel cells, beginning early in this century, is a fundamental step, the importance of which cannot be overestimated. There is still CO_2 produced in re-forming fossil fuel to hydrogen. However, because the fuel cell converts the chemical energy of the oxidation of hydrogen to water at twice the efficiency of the oxidation process for gasoline (which drives a combustion heat engine subject to the Carnot limitation on energy conversion), the amount of CO_2 used in the re-forming will be about half that emitted to the atmosphere in the present system. It is important to understand the need to eliminate all massive sources of CO_2 injection into the atmosphere. If all vehicles were run on fuel cells using fossil fuel re-forming to produce the hydrogen, the greenhouse effect would be delayed and the various biosphere-threatening processes associated with it would be slowed down. Their *elimination* is essential. We must also plan to meet the great increases in world energy needs as China industrializes. There is also a threat to the stability of the climate because of the unexpectedly greater increase in other greenhouse gases, particularly methane; and this is due simply to the growth in factory farming and the increased

demand for protein that accompanies improved living standards and population growth.

There are two general ways in which a zero increase in atmosphere CO_2 may be achieved.

1. Eliminate the use of fossil fuels as sources of hydrogen. Use solar-powered (photovoltaic or photoelectrochemical) electrolysis of water (the solar-hydrogen economy). In calculating the economics of such processes, one has to allow for the reduction in the cost of environmental damage that would occur by replacing CO_2-producing energy schemes with those that do not release CO_2.

2. Methanol has advantages as a carrier of hydrogen because it is a liquid and stable at room temperature. Instead of manufacturing methanol from a fossil fuel such as methane or coal, it would be possible to extract CO_2 from the atmosphere (Section 15.3.2), produce H_2 from solar-based water decomposition, and combine H_2 and CO_2 to CH_3OH. The CO_2 used up to do this would be re-injected to the atmosphere again upon the re-forming of methanol to H_2 for the fuel cells, but there would be no net buildup of CO_2. Further, we would not have to worry about the exhaustion of fossil fuels, which, as far as oil at an acceptable price is concerned, will occur before 2040.

These are choices to be made in the coming decades, and they will be influenced largely by economics, but also by the political influence of the largest corporations. Either path would lead to the elimination of the CO_2 contribution to the greenhouse effect.[22]

13.13. FUEL CELLS: THE SUMMARY

While batteries and fuel cells used to be the subject of a chapter in electrochemical books, the decision of the Daimler-Benz company in 1997 to develop fuel cells for the electric drive in their cars brought the fuel cell into clear focus; the battery suddenly took second place to environmentally friendly cars. The fuel cell directly converts the energy of chemical reactions to electricity and without moving parts, in contrast to the two-stage method of our present way of obtaining electricity (heat to mechanical work and mechanical work to a generator). Batteries store electricity produced elsewhere, and make it instantly available when a circuit is closed. They have their own market independently of whether they will be used in any automotive applications.

The history of fuel cells is lengthy. The first fuel cell, indeed, was produced in 1839 by a British judge, Sir William Grove. It was not until 1959 that Tom Bacon, a member of the family of Francis T. Bacon (who first enunciated the scientific method of experimentation and communication) made practical a 5-kW fuel cell. Tom Bacon,

---

[22]Although, at the end of the twentieth century, the accident at Chernobyl has made the use of fission reactors (eventually breeders) politically unacceptable, it must be recalled that our society can be run on electricity from nuclear reactors, with hydrogen as the storage medium and fuel for transportation. Fail-safe reactor schemes have been described in the literature. The eventual choice between nuclear energy and renewables will be one of cost.

indeed, remained the dominant world figure in fuel cell development until his work was taken up by NASA and made the basis of the fuel cells used in space vehicles.

Electrochemical kinetics can be easily applied to fuel cells and allow the general outline of their functioning to be understood. However, the kinetics given in the mathematical equations in this chapter are greatly simplified, for it is assumed in them that the fuel cell consists simply of two planar electrodes and that it has none of the complexities of the porous electrodes that are essential in the working of a real fuel cell. With flat plate electrodes, the power density would be impractically low. Nevertheless, from these equations one can learn that the functioning of fuel cells stands upon two legs. One is oriented to electrode kinetics and is essentially a matter of the rate of the interfacial electron-transfer reactions on the electrode surfaces, a subject with all the intricacies and opportunities of theoretical electrode kinetics; while the other leg is a matter of transport of the fuel to the surface and is oriented toward questions of diffusion and ion transport involving the three-phase boundary between the solid conductor, the solution in contact with it, and the fuel in gaseous form.

Modern fuel cells can be divided into four types. The one that comes most clearly out of history and the work of Bacon is the one which today has the least forward motion: the alkaline fuel cell, which is oriented to running on pure hydrogen. The phosphoric acid fuel cell was developed because of the need to reduce the expensive electrocatalyst by achieving superior performance through raising the temperature to 200 °C (which is possible with 98% phorphoric acid as the electrolyte). Until the 1990s, this cell was leading in the degree of development. Two other cells are important, and they are the high-temperature cells; one uses molten carbonates in temperatures above 650 °C; the other uses a solid electrolyte, zirconia-yttria and conducts largely by means of the mobility of the oxide ion therein; it runs at around 1000 °C.

The solid polymer electrolyte fuel cell was invented by Grubb in 1957. Since the mid-1980s it has been viewed as the most suitable fuel cell to develop for automotive transportation, partly because the liquid solution it contains is shrouded in a membrane that is a proton conductor. Furthermore, it works in an acid environment and therefore the CO_2 taken in with oxygen forms no blacking carbonate, as in the alkaline fuel cell, but is simply rejected back to the atmosphere. The membrane used in the United States is called Nafion, and is a polymer in which the proton produced by the electrochemical oxidation of hydrogen is mobile. The solid polymer electrolyte fuel cell (a proton conductor) has features in common with the high-temperature oxide anion-conducting fuel cell mentioned above. At present, this type of fuel cell is more commonly referred to as the proton exchange membrane fuel cell (PEMFC).

The electrochemical engine is an important concept and was first suggested by Douglas Henderson at General Motors in 1967; it uses a combination of a fuel cell and an electric motor. An important question for the practical use of electrochemical engines is what kind of fuel is to be used. In 1991 Roger Billings suggested the on-board re-forming of natural gas to produce hydrogen to be led into the cells. However, it is claimed that GM had been working on the on-board re-forming of

methanol since the mid-1980s. This concept has now been applied by a number of automotive concerns, led by the Daimler-Benz in Germany, because no massive supply of hydrogen, the fuel cell fuel, exists at present. Instead, common hydrocarbon fuels are being used. Methanol is the easiest to re-form (at 250 °C), but gasoline itself can be re-formed—though at 800–1000 °C—and its availability makes it the fuel of choice. The re-formed hydrocarbon produces CO_2 (rejected to the atmosphere) and the H_2 for the fuel cells which drive the electric motors in a car. Much depends, therefore, upon the quality of the re-former and how cleanly it can produce hydrogen and CO_2 only ($CH_3OH + H_2O \rightarrow CO_2 + 3H_2$). The trouble is that it produces a little bit of CO as well, and this is a dangerous poison for the electrodes. Thus, the success of the re-forming depends on completely removing the CO from the gas stream before the hydrogen reaches the anode.

The commercialization of the proton exchange membrane fuel cell depends upon a number of steps, the first of which were first taken in universities. Bacon's realization of the first 5-kW cell depended on an electrochemist, Rex Watson, from the University of London. The mechanism of the reduction of oxygen in acid solution, and the emphasis on Temkin kinetics, came out of work carried out by Damjanovic and Brusic at the University of Pennsylvania in 1967. The low Tafel slope to which this mechanism gives rise greatly reduces the polarization losses at the cathode and is key to understanding the working of oxygen reduction in acid-containing fuel cells. Cahan was responsible for the development of much helpful theory concerned with the distribution of current near the three-phase boundary in the pores of cells (a concept originated by Carl Wagner in reaction to a question put to him by early fuel cell scientists at the Pratt Whitney company). The development from the Cahan calculations of the mathematics for this distribution led Srinivasan to write a fundamental paper on how these ideas would affect the design of fuel cells. Distribution of the catalyst (which Cahan's work has shown was only useful near the three-phase boundary) was the key to the higher use factor of platinum, which was implemented later in work led by Srinivasan at Los Alamos National Laboratory (compare Gottesfeldt, 1980s) and at Texas A&M University. Ballard Power Systems, Inc., was the site of magnificent progress in developing the proton exchange membrane fuel cells to the 50-kW level suitable for cars. This in turn was based on work at Texas A&M, which helped to achieve the 1 kW/kg for fuel cells that underlay the progress Ballard achieved. The membrane is still undergoing further research in Japan, the United States, and Germany.

In spite of all these triumphs and the Daimler-Benz effort (and its effect upon all the other automotive makers), the fuel cell may need a boost for starting and acceleration when it is used to power vehicles. However, the fuel cell has a higher energy-to-weight ratio than any battery. Any short-term boost in power that is needed can be given by an auxiliary battery that can be recharged from regenerative energy braking, or indeed from the fuel cell under cruising conditions. An alternative would be to use electrochemical condensers (Section 13.19) as the boosting power source.

What does all this mean in the larger context? The plain fact is that for more than two centuries we have based the development of our economy on heat engines, although Carnot showed in 1824 that such engines cannot produce mechanical power at greater than a certain efficiency, which under practical circumstances comes to somewhere between 20 and 30% for internal combustion engines. Electrochemical engines are free from such limitations. They have theoretical maximum efficiencies in the 90% region, and their practical efficiencies lie between 50 and 65%, a doubling of the energy conversion efficiency obtained from combustion motors. The advantage is more than that, for they eliminate the pollution created by incompletely combusted hydrocarbons which, after 30 years of attempts in Detroit to change this situation, still are emitted by cars working on the internal combustion principle. The electrochemical fuel cell is not an evolutionary development. It will cause, and it has begun to cause, a revolution in technology that will affect virtually all aspects of our lives.

It is suggested in this book that there is a second fuel cell principle. The first one is concerned with the highly efficient conversion of chemical fuels to electrical energy. The second is concerned with direct electrochemical production of materials without the use of outside electricity. In fact, the method would produce electricity as a by-product that can be transmitted to places far away.

We now have plans (which are very likely to be realized) that put us halfway through the change we need to go from our pollutive, indirect method of producing energy to the direct method of converting energy from light and perhaps from new methods of nuclear reactions which will produce electricity and hydrogen, and allow us to develop our technologies in a steady-state nonpollutive future.[23]

It should be noted here in passing, for they are present only at whisper strength, that there are new ideas about energy so novel and so different that they are too fresh to be explained in a textbook about electrochemistry. They concern the startling concept that one might be able to use the so-called "energy of the vacuum," which (it is claimed) contains within it vast amounts of energy from the surrounding electromagnetic radiation of the entire universe. However, even if this book is used for as long as its predecessor, the practical implementation of such ideas surely lies in a time when this book will have ceased to be of use.

Further Reading

Seminal

1. W. Grove, *Phil. Mag.* **14:** 127 (1839). The first recognized paper on fuel cells.
2. L. L. Mond and D. Langer, *Proc. Roy. Soc. London* **A46:** 296 (1989). Practical development of Grove's work.

---

[23]Toward the end of the century there were indeed rumors about "the future of energy." There is no doubt that just ahead lies the use of various forms of renewable energies, and perhaps new forms of nuclear power that operate on different principles than those of the past.

3. F. W. Ostwald, *J. Electrochem.* **1:** 122 (1894). Presidential address to Bunsen Gesellschaft; predicted pollution if heat engines are used for energy production. Recommended electrochemical pathway.

4. W. Jacques, *Harper Mag.* **94:** 199 (Dec. 1896, 1897). Worked out in detail the reduction of fuel costs for a boat to cross the Atlantic using electrochemical or chemical engines.

5. E. Bauer, W. D. Treadwell, and G. Trumpler, *Z. Elektrochem.* **27:** 199 (1921). First carbonate fuel cell.

6. E. C. Potter on F. T. Bacon, in *Trends in Electrochemistry.* J. O'M. Bockris, D. A. J. Rand, and B. J. Welch, eds., 7.1, Plenum, New York (1977). A summary of Bacon's work.

7. E. Justi and J. Winsel, *Cold Combustion,* Verlag, Chemie, Wiesbaden, Germany (1962). A general account of fuel cells before the 1960s.

8. J. O'M. Bockris and S. Srinivasan, *The Electrochemistry of Fuel Cells*, McGraw-Hill, New York (1969). Electrochemistry and electrochemical engineering associated with practically all types of electrochemical energy conversion systems, as well as their applications. An advanced presentation still relevant in the 1990s.

9. F. T. Bacon, Fuel Cells, in *Trends in Electrochemistry,* J. O'M. Bockris, D. Rand, and B. Welch, eds., Plenum, New York (1976). An address to the meeting of the Ivth Australian Conference in Electrochemistry.

10. A. Damjanovic and V. Brusic, *Electrochim. Acta* **13:** 615 (1967). The basic paper deducing $O_2 + H^+ + e \rightarrow OH_{ads} + O_{ads}$ as the rds for O_2 reduction in acid solution.

11. E. Yeager, J. D. E. McIntyre, and M. J. Weaver, Electrocatalysis, *Proc. Electrochem. Soc.* **84-12:** 247 (1984).

12. A. Damjanovic and M. A. Genshaw, *Electrochim. Acta* **15:** 1281 (1970). Experimental proof of Temkin behavior. O adsorption linear with potential.

13. H. Wroblowa, M. L. B. Rao, A. Damjanovic, and J. O'M. Bockris, *J. Electroanal. Chem.* **15:** 139 (1967). The stationary potential of Pt contact with O_2 in solution.

14. A. J. Appleby, in *Comprehensive Treatise of Electrochemistry,* B. E. Conway, J. O'M. Bockris, S. U. M. Khan, and R. E. White, eds., Vol. 7, p. 173, Plenum, New York (1983). Electrocatalysis of O_2 reduction.

15. P. Zelenay, B. R. Scharifker, J. O'M. Bockris, and D. Gervasio, *J. Electrochem. Soc.* **133:** 2262 (1986). The small adsorption of sulfonic acids on Pt.

16. M. A. Enayetullah, T. D. deVilbiss, and J. O'M. Bockris, *J. Electrochem. Soc.* **136:** 3369 (1989). The high solubility of O_2 in trifluoromethane sulfonic acid.

Modern

1. A. Parthasarathy, S. Srinivasan, and A. J. Appleby,*J. Electroanal. Chem.* **339:** 101 (1992). Reduction of O_2 at C-supported Pt microcrystal/Nafion interfaces.

2. A. J. Appleby, *Energy* **21:** 145 (1996). A review (640 refs).

3. S. Chalk, *Fuel Cells for Transportation,* U.S. Department of Energy, Washington, DC (1997). A brief review, many figures.

4. E. A. Ticianelli, J. G. Berry, and S. Srinivasan, *J. Appl. Electrochem.* **21:** 597 (1991). Attaining small particle size for Pt in fuel cells.

5. J. B. Goodenough, A. Hamnet, B. J. Kennedy, and S. A. Weeks, *Electrochim. Acta* **32:** 1233 (1987). XPS studies of platinized carbon.

6. K. Machido and M. Enyo, Tungsten bronze electrodes doped with platinum (methanol oxidation), *J. Electrochem. Soc.* **135:** 1955 (1988).

7. K. Wang, H. A. Gasteiger, N. M. Markovic, and P. V. Ross, *Electrochim. Acta* **16:** (1996). Methanol oxidation on Pt-Sn and Pb-Rh.

8. A. V. Tripkovic and K. Popovic, *Electrochim. Acta* **41:** 2385 (1996). Oxidation of methanol on Pt (110) single crystals.

9. H. Gasteiger, N. Markovic, P. N. Ross, and E. J. Cairns, *J. Electrochem. Soc.* **141:** 1296 (1994). Temperature-dependent methanol electro-oxidation on well-characterized Pt-Ru alloys.

10. A. K. Schukla, M. K. Ravikumar, A. S. Arico, C. Candiano, V. Antonucci, and N. Giordano, *J. Appl. Electrochem.* **25:** 528 (1995). Methanol oxidation on Pt-WO_3.

11. T. E. Springer, M. S. Wilson, and S. Gottesfeld, *J. Electrochem. Soc.* **140:** 3513 (1993). Modeling in polymer electrolyte fuel cells.

12. S. Chalk, J. F. Miller, and S. R. Venkataswaren, paper presented at Fifth Grove Fuel Cell Symposium, London, 1997. A review of the fuel cell programs of the U.S. Dept. of Energy.

13. Allison Gas Turbine Division, Research and Development of Proton Exchange Membrane Fuel Cells for Transportation, U.S. Department of Energy, Office of Transportation Techniques, Washington, DC, 1996. (Available through National Technical Information Service, Springfield, VA.)

14. M. S. Wilson and S. Gottesfeld, *J. Appl. Electrochem.* **22:** 1 (1992). Making fuel cell electrodes with small Pt loading.

15. A. J. Appleby and F. R. Foulkes, *Fuel Cell Handbook,* Van Nostrand, New York (1989).

16. S. Srinivasan, O. A. Velev, A. Parthasasathy, and D. J. Manko, *J. Power Sources* **36:** 299 (1991). O_2 reduction in PEM cells.

17. A. John Appleby, J. O'M. Bockris, E. B. Yeager, T. Robert Selman and J. T. Brown, "Penner Report on Fuel Cells," U.S. Department of Energy, Washington, DC (1986).

18. S. Srinivasan and C. Lamy, "The Direct Methanol Fueled Fuel Cell," in *Modern Aspects of Electrochemistry,* B. E. Conway, R. E. White, and J. O'M. Bockris, eds., Vol. 34, p. 1, Kluwer Academic, Plenum, New York (1999).

19. J. M. Gür and R. A. Huggins, *J. Electrochem. Soc.* **139:** 295 (1992). Carbon to electrical energy in a fuel cell.

13.14. ELECTROCHEMICAL ENERGY STORAGE

13.14.1. Introduction

The first part of this chapter has been concerned with the fact that some chemical reactions that take place spontaneously can be split into two electrode reactions which, when joined together in an electrochemical cell, give rise to electric power (electrochemical energy conversion—fuel cells). We now turn to processes by which an

outside source of electricity can be used to drive a chemical reaction in a cell *up* a free energy gradient, the reaction being split up to occur in parts, one on each electrode. When the electricity to be stored has brought about these two electrode reactions, the newly formed substances, which are ready to react together *down* the free energy gradient, when electronically joined together are equivalent to stored electrical energy. Since they were manufactured by pushing the chemical reaction up the free energy hill, the two electrode processes will work together to *discharge* the substances on the electrodes back to the state they were in before "charging." In this discharging, the overall reaction runs down the free energy hill, producing electricity in an outside circuit containing a "load," which can be anything from a bulb in a flashlight to the electric motor running a car.

One electrochemical electricity storer known to all is the lead-acid battery, used to start up internal combustion engines and for numerous other purposes. But the virtues of this particular storer (a high power-to-weight ratio; cheap materials; facilities for manufacture throughout the world) are associated with the vice of heaviness (watt hours per kilogram of only 25–40). Because it is this rechargeable battery that is the most visible to the public, the image it has engendered is that "batteries are heavy." In fact, numerous newer rechargeable batteries have been demonstrated (and several are in commercial production) that are up to about five times lighter per unit of energy they store than the venerable lead-acid cell.

Batteries are grouped into two divisions, which have names based more on history than on rationality. The ones that cannot be recharged are called *primary.* Those that can be recharged are called *secondary.* Secondary batteries are, of course, more useful than those illogically called "primary," though it may be that in actual numbers sold, the primary cells (e.g., $Zn\text{-}MnO_2$) are far in excess[24] (see Table 13.5).

There is no question of a "best" battery, because batteries are used for so many different purposes, each with its different requirement. For example, batteries for hearing aids must be above all tiny. Batteries for pacemakers must be above all reliable. Batteries for torpedoes must be stable during storage ("have a good shelf life") and give high power for short times. And batteries for driving submarines while submerged must be very large and certainly rechargeable. The market for batteries is both huge and world wide (see Tables 13.5, 13.6, and 13.7).

The world of batteries is a much larger world than that of fuel cells. Even if, as seems probable, a generation from now much of our conversion of fuels to electricity would occur via fuel cells, the need for batteries will not diminish. However, it is worth recording a few words here about the outlook at the century's end as to the possibilities for batteries as automotive power sources. There seems little doubt that for long-range land and sea transportation, the future lies with the fuel cell, not only because of the reduction of air pollution, but also because of the greater range made possible and the absence of a long wait for recharging to occur. However, for the small "shopper" or

---

[24]The nonrechargeable version of the $Zn\text{-}MnO_2$ cell is used for flashlights, tape recorders, computers, portable TVs, etc.

TABLE 13.5
Total Primary and Secondary Battery Market Forecast (billions, 10^9)

	1995	1997	2001
Primary batteries[a]	27 (93%)	35 (93%)	46 (91%)
Rechargeable batteries[b]	2 (7%)	3 (7%)	5 (9%)
Total	29 billion	38 billion	51 billion

Source: Reprinted from K. Kordesch and Faistaucer, "RAM Cells with Low-Cost Chargers May Compete with MNO_2-ZN Primaries on the Global Consumer Market" in *Batteries for Portable Applications and Electric Vehicles*, Holmes and Landgrebe, eds., Electrochemical Society, 1997, p. 924–925. Reproduced with permission of the Electrochemical Society.
[a] Zn-carbon plus Alkaline.
[b] Without large format and SLI batteries.

commuter car (80% of all car use!) there may be some use for batteries as automotive power sources. The facilities for overnight, easy recharge are available in every home. Battery-driven cars would emit no CO_2 or unsaturated smog-causing hydrocarbons, though they would lead to an increase in SO_2, NO_x, and CO_2 at coal-burning (Carnot limited) electricity plants.

TABLE 13.6
Geographic Sampling of Alkaline Cell Penetration

Region or Country	Population (millions)	$ Per Capita Income (10^3)	No. of Batteries per Capita	No. of Batteries (millions)	No. of Alkaline Penetration
U.S.A.	260	24.8	10.9	2,843	86.5
Canada	29	13.2	5.5	158	74.7
U.K.	58	17.8	6.8	395	62.0
Germany	81	22.9	5.9	482	51.7
Italy	57	20.5	6.6	377	59.4
Spain	39	14.0	7.4	292	39.4
France	58	22.3	8.9	515	58.4
Russia	149	2.7	3.5	525	26
India	870	0.23	1.5	1,300	0.3
Middle-East[a]	110	3.6	3.6	401	14.3
China	1,200	0.43	3.1	3,750	3.2
S. America[b]	348	3.4	5.4	1,890	30.2
Total				13 Billion (10^9)	

Source: Reprinted from K. Kordesch and Faistaucer, "RAM Cells with Low-Cost Chargers May Compete with MNO_2-ZN Primaries on the Global Consumer Market" in *Batteries for Portable Applications and Electric Vehicles*, Holmes and Landgrebe, eds., Electrochemical Society, 1997, p. 924–925. Reproduced with permission of the Electrochemical Society.
[a] Egypt, Israel, Morocco, South Arabia.
[b] Argentina, Brazil, Chile, Colombia, Mexico, Venezuela.

TABLE 13.7
World-Wide Use of Consumer-Type (Small format) Rechargeable Batteries
(millions, 10^6)

	1995	1997	2001
Nickel cadmium	1,590 (81%)	1,710 (65%)	1,940 (42%)
Nickel metal hydride	310 (16%)	600 (23%)	1,716 (37%)
Lithium batteries	30 (2%)	250 (9%)	700 (15%)
Rechargeable alkaline	20 (.1%)	70 (3%)	284 (6%)
Total	1.95 million	2.63 million	4.64 million

Source: Reprinted from K. Kordesch and Faistaucer, "RAM Cells with Low-Cost Chargers May Compete with MNO_2-ZN Primaries on the Global Consumer Market" in *Batteries for Portable Applications and Electric Vehicles*, Holmes and Landgrebe, eds., Electrochemical Society, 1997, p. 924–925. Reproduced with permission of the Electrochemical Society.

13.15. A FEW HIGHLIGHTS IN THE DEVELOPMENT OF BATTERIES

13.15.1. History

There is archeological evidence (König, 1936) that a battery and an electrochemical gold plating device existed about 2000 years ago in the country now known as Iraq. The battery seems to have been an iron-copper device, but no residues of an electrolyte could be identified. It was Volta (along with Galvani, generally credited with the founding of electrochemistry) who, in 1800, reported on the electricity produced in his experiments in placing silver and zinc together but separated by absorbent paper saturated with electrolyte. Volta's signal achievement was followed by that of Daniel (1836), who presented a cell consisting of a copper electrode in a copper sulfate solution, accompanied by a zinc electrode in a zinc sulfate solution, the two solutions being separated by a porous membrane. PbO_2, to be used in the lead-acid battery, and MnO_2, to be a part of the most used primary cell, had both been introduced as battery materials by 1866.

Grove's fuel cell of 1839 may have had a trigger effect on Planté who announced the "lead-acid accumulator", the first rechargeable or secondary battery, in the same year. A whole generation grew up before the second rechargeable cell came into being, the nickel-cadmium cell of 1896, its selling point being a longer life than that of the lead-acid cell. Finally, among the "historical" batteries came Edison's contribution, the nickel-iron battery, used now for situations (as in power for the lighting sets on trains) where lengthy periods in the uncharged state are common.

The twentieth century witnessed a proliferation of new primary cells and a few new rechargeable ones. In this book, we choose to outline only six of these and to stress the newer and underplay the older.

13.16. PROPERTIES OF ELECTROCHEMICAL ENERGY STORERS

13.16.1. The Discharge Plot

The discharge plot is a plot of the potential available from a cell versus the time during which it discharges. In the following diagrams, ideal and real plots are given. The ideal plot is based on the assumption that the electrode and reactions each occur at the same steady rate until the material on the plates, placed there in the charging process, is exhausted. Then the potential drops to zero (see Fig. 13.36).

The real situation is less stark than that indicated. For one thing, the material on the plates may not be uniform and that which electrochemically converts to another substance at first may do so more easily than the material at deeper layers on the electrodes. One of the properties by which the value of a battery is judged is the length of the almost flat plateau, i.e., a good battery discharges for many hours with only a
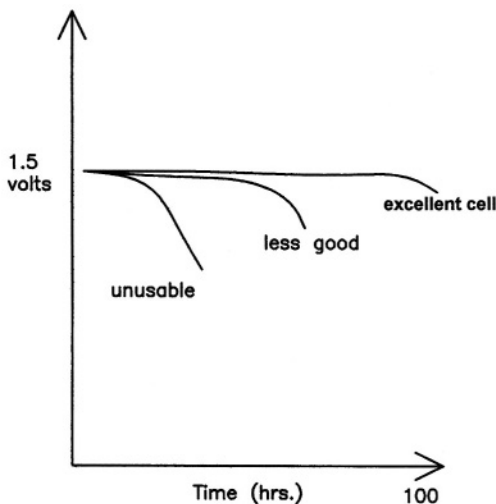
Fig. 13.36. A schematic of the "discharge curve" for batteries. In this ideal model the potential during discharge through a given load remains constant until the active material on the battery plates has all been used up, whereupon the volts would suddenly become zero. An excellent performance of a real cell is shown in the top line of the figure. Some systems considered as useful for batteries suffer a rapid fall in potential, as shown in the lowest curve.

small decline in potential. Less good batteries "slope off" with time, and become in practice less usable (Fig. 13.37).

13.16.2. The Ragone Plot

The Ragone plot is a plot of the power per unit weight ("specific power") that can be pulled from a battery against the energy per unit weight (specific energy) that it will deliver. The first is measured in watts per kilogram and the second in watt hours per kilogram.

Figures 13.37(a) and 13.37(b) show an idealized Ragone plot together with Ragone plots for several batteries. It can be seen that one cannot say that a certain battery has a definite specific power and energy. Rather, each of the batteries shown in Fig. 13.37(b) has a characteristic range of values of the energy it can store (watt hours per kilogram) if a certain power were constantly demanded of it (watts per kilogram). If one needs to use it at a high power level, the battery will have less energy available. If one wants the battery to give back a maximum amount of energy, one had better ask for it at a low power level.[25]

Here one meets a paradoxical generalization. Batteries (the electrochemical energy storers) are good at providing high *power* levels, but the amount of energy they can store per unit weight is not so great (see the following figures) because they can at best use up all the material on their plates. Fuel cells, on the other hand, can provide a much larger energy per unit weight (for they simply convert whatever amount of chemical fuel is provided them), but their power per unit weight tends to be more limited because of the low i_0 of the cathodic oxygen reduction reaction.[26]

This hints, then, at a splendid symbiosis between batteries and fuel cells. In automotive transportation, short bursts of high power are needed for starting and in periods of acceleration, e.g., in passing. One way[27] of providing this is to have lightweight batteries of low capacity (i.e., not containing much material on the plates) for the periods when the power thrusts are needed (see also the role of electrochemical capacitors, Section 13.19), and use fuel cells for the normal steady running. The fuel cell can continue to power a vehicle for a distance that depends simply on the amount of fuel in its tank.

---

[25]Of course, this is similar in trend to the performance of an internal combustion engine; the faster one drives (i.e., the more power demanded), the fewer miles per gallon one obtains from the fuel (i.e., the less is the energy available from a given amount of stored fuel).

[26]However, the superiority of batteries over fuel cells in respect to specific power is gradually being ceded to fuel cells. A well-catalyzed proton exchange membrane fuel cell can reach up to 1 kW kg^{-1}. Only batteries involving high temperatures can do as well.

[27]There are several other ways of providing "hybrids," cars that are powered by two different power sources that can be switched on to suit needs. Compressed air systems provide a pollution-free way of making a boost available when starting.
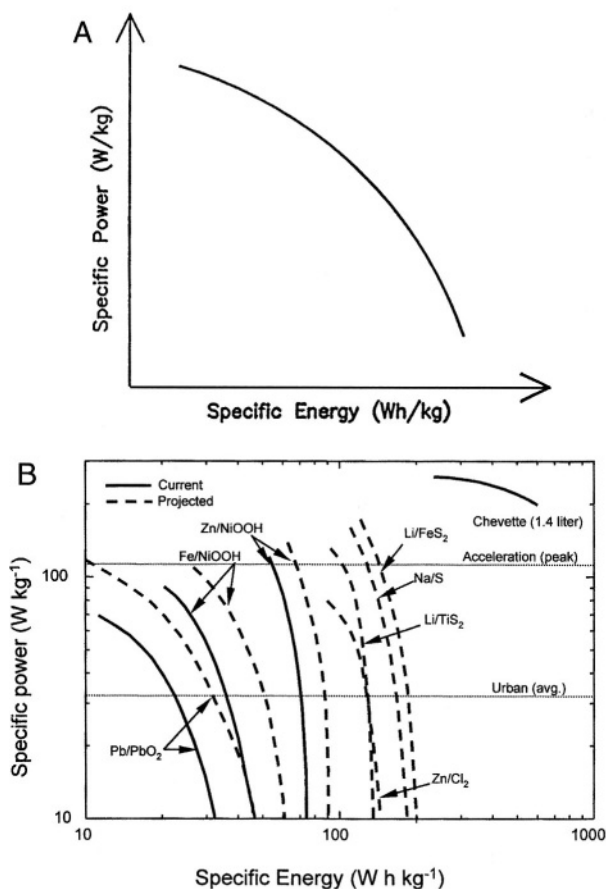
Fig. 13.37. (A) The idealized Ragone plot. (B) Real Ragone plots for certain batteries. The specific energy declines with the increase of the specific power demanded.

13.16.3.   Measures of Battery Performance

The most important property of a battery is undoubtedly its energy density, i.e., the watt hours per kilogram. Batteries are storers and this figure tells how much they store for a unit of weight. A similar measure, less used, is watt hours/liter. As explained above (Ragone plot), any energy density figure given for a battery type depends on the rate of discharge. Nevertheless, most battery types quoted in the literature are described with certain figures for their energy densities and this is then the watt hours kg^{-1} for the typical rates of discharge at which the particular type of battery is used.

TABLE 13.8
Some Batteries: Their Status Toward 2000

	Availability	Cycles	Watt hours
Lead-acid	Yes	300–400	35–40 kg
Nickel cadmium	Yes	700–1200	45–55
Zing manganese dioxide	Yes	25	8–64
Zinc-air	Yes	Mechanically recharged	200
Nickel metal hydride	Yes	700–1200	150–200
Lithium ion or Lithium polymer	Probable	400–1200	100–200
Aluminum-air	Not yet	Recharged	500

Thirty to three hundred watt hours kg^{-1} is the range of figures for currently available battery types. The watt hours kg^{-1} range is 100–250. These figures are 3–4 times smaller than one can calculate by making zeroth approximation assumptions, e.g., by neglecting the weight of the container, the plates to which the active material is attached, the overpotential, and the wasteful IR drop in the solution between the electrodes. The next most quoted figure about a battery (but again, please look back at the Ragone plots) is the power density, watts kg^{-1}. This varies over a greater range (50–1000 Watts kg^{-1}) with the various available battery types, and the maximum figures quoted will not be those often used because of the loss of energy storage capacity that goes with working at high rates.

Several other characteristics must be taken into account in considering whether a given type of battery (Table 13.8) is viable for a certain application. The number of times it can be recharged is an important criterion, but to make the figure meaningful, one must state the depth to which the cell is being discharged (usually 90%). For most of the present-day applications, rechargeable batteries have to be able to be "cycled" only about 100 times. However, for applications such as energy storage by electric utility power plants for electric vehicles, the required number of cycles is in the region of 1000. Few batteries continue to be useful beyond ~1000 recharges at high depth of discharge. However, in some systems 50,000 recharges are possible at a 40% depth of discharge (e.g., nickel-hydrogen batteries in low earth-orbiting satellites).

The calendar life of a battery is greatly dependent on the mode of use. It is also affected by competing degradative side reactions. For example, batteries involving Zn tend to form spikey dendritic growths upon repeated high current density recharges, and these may eventually pierce the separator and "kill" the cell by creating a short circuit between cathode and anode.[28]

---

[28] The short-circuiting causes extremely high currents to pass until the separator is burned, etc.

Another kind of life is the shelf life, how long a battery remains viable if it is left unused in the charged state. This varies very much from one battery type to another, but nearly all batteries deteriorate under such conditions because of slow corrosion reactions that change the nature of the electrodes.

13.16.4. Charging and Discharging a Battery

Charging a battery takes more energy than one gets out on discharging it. The reasons are not difficult to see. In an idealized zeroth approximation, the two amounts would be the same. However, this would imply 100% efficiency in all aspects of the electrochemical processes of charge and discharge. Thus the Butler–Volmer equation at once tells one that at zero overpotential (i.e., zero energy loss), there is a zero reaction rate. Hence, when a cell has an open-circuit potential measured against a voltmeter having, e.g., a 10^{12}-ohm resistor, there will be registered (for virtually zero current drawn) effectively the maximum potential of which the cell is capable. Charging the cell means overcoming nature, i.e., going against the free energy gradient and making the cell reaction "go backward." Hence, the necessary electrical energy applied to the cell will have to be greater than the maximum possible (thermodynamically reversible) potential it could exhibit. One will in fact have to apply not only this potential, but also whatever overpotentials the electrode reactions need to function at the rate of charge one chooses. Of course, there will be an IR loss through the cell, also, and in batteries this is often important and even dominating. Hence, on charging, the needed potential is

$$E_{charge} = E_{rev} + \Sigma |\eta| + IR$$

where the Σ indicates the sum of all overpotentials taken with a positive sign.

Going downhill, i.e., discharging the cell, is subject to similar thinking but of course reversed. The potential available at the terminals of a battery in discharge will be the thermodynamically reversible potential but now *diminished* by the sum of the overpotential and the IR drop.

$$E_{discharge} = E_{rev} - \Sigma |\eta| - IR$$

Thus, only at extremely low current drains (e.g., at nanoamperes cm^{-2}) do E_{charge} and $E_{discharge}$ come near each other. All this can be put together in a scheme that looks like that in Fig. 13.38.

13.17. SOME INDIVIDUAL BATTERIES

13.17.1. Introduction

The mode of presentation here will be to describe the more important batteries of the late 1990s according to three divisions. First, one sees three batteries (lead-acid, nickel-cadmium, and zinc-manganese dioxide) which, in various forms, have been in
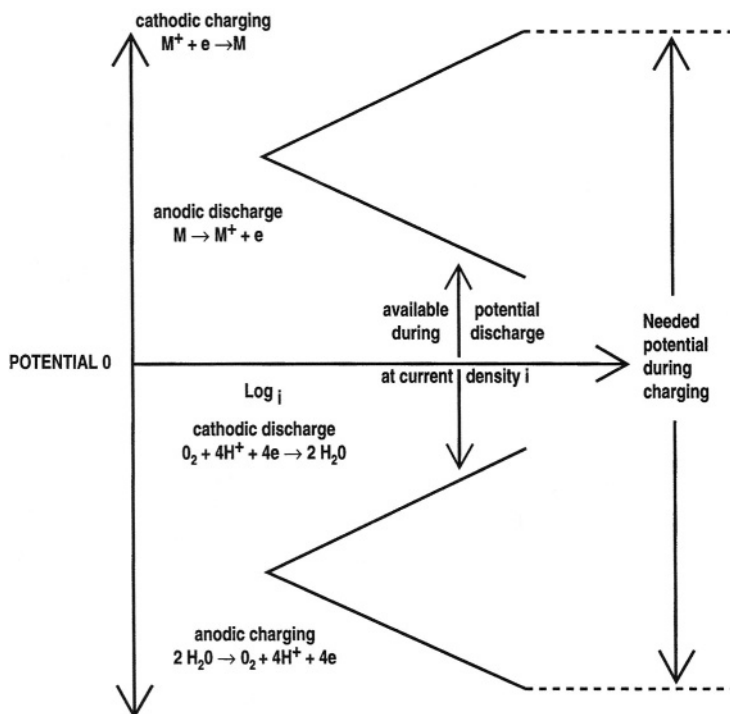
Fig. 13.38. Potential log current density relations for charge and discharge of a battery. Potential axis is not drawn to scale.

use for many decades ("classical storers"). Next, come three batteries at the forefront of modern development and in the early stages of commercialization (zinc-air, metal hydride, and lithium batteries). These metal hydride and lithium ion batteries are being commercialized for cellular phones and laptop computers. Finally, a battery will be briefly discussed that is a "maybe" for the near future.

13.17.2. Classical Batteries

  13.17.2.1. *Lead-Acid.*   Lead-acid batteries stem from the time of Planté (1860).

During discharge at the anode (negative terminal) the main reaction is

$$Pb = HSO_4^- \rightarrow PbSO_4 + H^+ + 2e_0^-$$
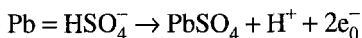
At the cathode (the positive terminal during discharge) the corresponding reaction is

$$PbO_2 + 3H^+ + HSO_4^- + 2e \rightarrow PbSO_4 + 2H_2O$$
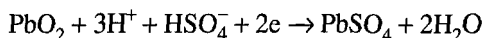
The overall cell reaction during the functioning of the two electrode reactions to give back the stored electrical energy is, then, the sum of these two reactions. Using tables of free energy data, one finds that the corresponding $\Delta G°$ is = 372.5 kJ mol^{-1}. Using the well-known relation between the standard potential and the standard free energy change in the overall reaction in a cell, $nFE° = -\Delta G°$, one finds

$$E° = \frac{372.5 \times 10^3}{2.96493} = 1.93$$

Here, the 2 comes from the two electrons needed in the functioning of electrode reactions as formulated above. The factor 96493 is the faraday in coulombs mol^{-1}.

The lead-acid battery has a huge market as the starter battery for internal combustion engines. Here, a short pulse at high power, using up little energy, is all that is needed. This lead-acid cell performs very adequately for this purpose. Its poor energy density (only some 28–35 in watthours kg^{-1}) is of no consequence because the *energy* needed for each use is small (in normal use, the battery is promptly recharged after every use from the car's generator). This ideal niche, fitting the properties of lead acid (high power, low energy, density) and having a tremendous market, has led to a large world-wide manufacturing activity for this one kind of battery. During the 1960s and 1970s, it was effectively the only battery considered for electric cars, and the only one readily available at an acceptable cost. However, the lead-acid battery deteriorates because the discharge reactions form lead sulfate and the charging reactions do not completely reverse this formation [side reactions of hydrogen and oxygen evolution; "gassing"] so that lead sulfate gradually builds up on the plates). The life of a lead-acid battery, even under the easy conditions of a starter battery in an automobile (no deep discharges), seldom exceeds 4 years.

There is mild concern about the use of lead. Junked lead-acid batteries lying in waste dumps are occasionally said to introduce lead (a toxic metal) into the food chain.[29] Ingested, lead remains in the brain. (Hence the banning of Pb-containing additives in gasoline.) However, most of the lead in lead-acid batteries is recycled.

13.17.2.2. *Nickel-Cadmium.*

This cell was developed by Jungner in 1890. It was put forward as an improvement on lead acid for the energy density (50 W hr kg^{-1}) is about 30% greater (though the power density somewhat less) than that of lead acid. During the anodic discharge reaction, the negative anode is the site of the reaction:

$$Cd + 2OH^- \rightarrow Cd(OH)_2 + 2e$$
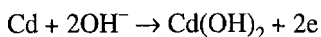
At the positive cathode there occurs

---

[29] But the transfer rate is low. Does it depend on animals ingesting battery parts? Could cows ingest lead in this way?

$$2 \text{ NiOOH} + 2\text{H}_2\text{O} + 2e_0^- \to 2 \text{ Ni(OH)}_2 + 2\text{OH}^-$$
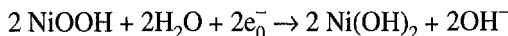
The thermodynamically reversible potential (zero current) for the overall cell reaction is 1.299 V (compare 1.93 V for the lead-acid cell). The greater potential of the latter means less cells in series for a given potential, e.g., if a 200-V motor is to be powered. A big advantage of the Ni-Cd rechargeable battery is not only the higher energy density but also the longer life. For a 25% depth of discharge (as is often the real situation), the cell can be recharged several thousand times. Now, if the battery is discharged to 75% of the full amount, it still accepts several hundred recharges, in this respect being significantly better than lead acid. The degradative processes in the Ni-Cd battery are evidently not very active. They involve slow changes in the NiOOH, which forms 2NiO_2–NiOOH. Some $[\text{Cd(OH)}_4]^{2-}$ is formed and tends to block pores in the electrode. The NiCd battery has been *the* rechargeable battery and has been used for diverse purposes in toys, electronic devices, and computers, etc. However, enthusiasm for this battery has been quelled if not quenched by the finding that Cd is a toxic metal. Thus, its presence in the above devices (near children in toys) poses a greater threat to the user than does the lead in car batteries, which is remote from children. An explicit diagram of this battery is given in Fig. 13.39.

13.17.2.3. *Zinc-Manganese Dioxide.* In 1866 Leclanché invented a galvanic cell in which the reduction of MnO_2 is the cathodic reaction in the cell's discharge. The corresponding anodic dissolution reaction is the oxidation of zinc. The Leclanche cell is a (so-called) dry cell, i.e., the ammonium electrolyte is immobilized in the form of a paste. There are three forms of the zinc-manganese dioxide batteries:
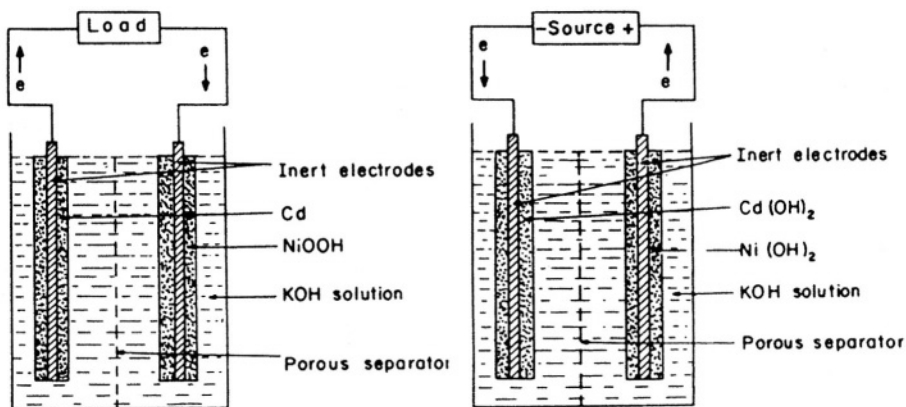
Fig. 13.39. (a) A nickel-cadmium cell at the beginning of discharging. At the cathode, $2\text{NiOOh} + 2\text{H}_2\text{0} + 2e \to 2\text{Ni(OH)}_2 + 2\text{OH}^-$; at the anode, $\text{Cd} + 2\text{OH}^- \to \text{Cd(OH)}_2 + 2e$. (b) A nickel-cadmium cell at the beginning of charging. At the cathode, $\text{Cd(OH)}_2 + 2e \to \text{Cd} + 2\text{OH}^-$; at the anode, $2\text{Ni(OH)}_2 + 2\text{OH}^- \to 2\text{NiOOH} + 2\text{H}_2\text{O} + 2e$.

1. The Leclanché form, in which the electrolyte is an aqueous solution of ammonium chloride supported in paste form.
2. A form of the Leclanché cell for extra-heavy duty in which the electrolyte is zinc chloride mixed with a minor fraction of ammonium chloride.
3. The alkaline-zinc manganese dioxide battery.

A schematic diagram of Leclanché's cell is shown in Fig. 13.40.

The electrochemistry of the $Zn-MnO_2$ battery is substantially more complicated than that of the lead-acid or nickel-cadmium batteries already described. The three forms of it are described above, although the focus in this book will be on the modern alkaline $Zn-MnO_2$ cell. Then, there are many allotropes of MnO_2 which occur in nature and each has different electrochemical properties (although the one most used is synthetically made). A further complication arises as to rechargeability. This cell has (during its long lifetime) been regarded as a "primary," i.e., a battery which, once manufactured, could be discharged but was not amenable to being recharged (a "throwaway"). However, as will be described, to a degree, this is no longer the case. Finally, it is not possible (as it has been with the two previous batteries) to write the cathodic reaction on discharge (the reduction of MnO_2) in a simple, unambiguous way. One might wonder, then, at the attention paid here to a battery that originated in 1866. However, the manufacture of manganese dioxide for use in batteries amounts to about 500,000 tons per year, and batteries based on MnO_2 are the most used of all.

On the other hand, there is no difficulty in writing an equation for the anodic reaction on discharge. It corresponds, as stated, to the oxidation of zinc, and can be written for an alkaline solution as:
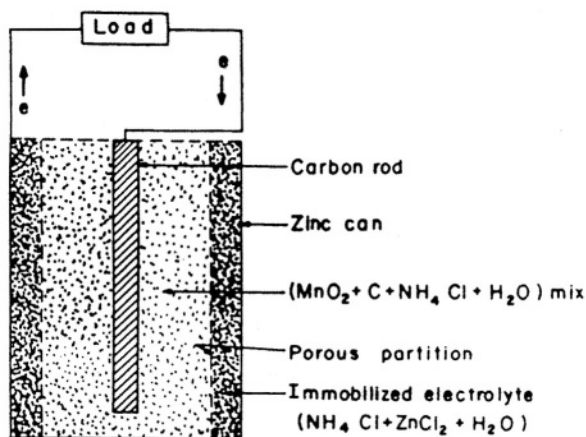
Fig. 13.40. Schematic diagram of a Leclanché cell.

$$Zn + 2OH^- \rightarrow ZnO + H_2O + 2e_0^-$$
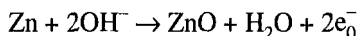
As to the cathodic reaction in discharge, it is the reduction of MnO_2, that has attracted considerable discussion. Thus, *if* the MnO_2 undergoes complete discharge in alkaline solution, the reaction can be written:

$$MnO_2 + 2H_2O + 2e \rightarrow Mn(OH)_2 + 2OH^-$$
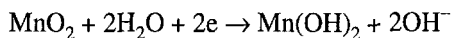
This is logically called "the two-electron reaction" and is equivalent to the reduction of the state of Mn from Mn^{4+} to Mn^{2+}. Viewing the discharge curve for MnO_2, one can see there are two fairly distinct sections (Fig. 13.41).

In the first stage, the potential varies substantially about +0.2 to –0.5 vs. a reference electrode of Hg-HgO. However, thereafter, the curve flattens out and the potential remains constant if the current density is low (e.g., 0.01 mA cm^{-2}) until the reduction reaction to $Mn(OH)_2$ has completed itself throughout the whole oxide mass
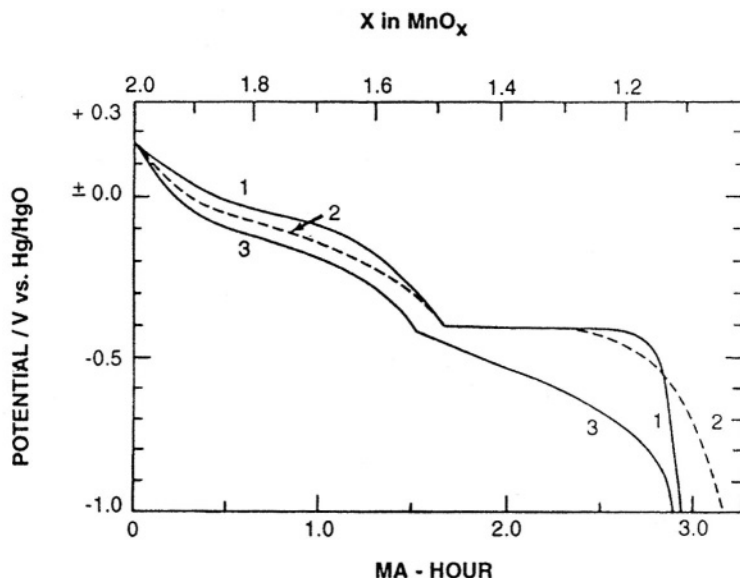
Fig. 13.41. Discharge curves for thin films of manganese dioxide on a graphite substrate in 9 *M* KOH at 23 °C. 1, i = 0.01 mA/cm^2; 2, i = 0.03 mA/cm^2; 3, i = 0.28 mA/cm^2. (Reprinted from T. N. Anderson, "The Manganese Dioxide Electrode in Aqueous Solution," in *Modern Aspects of Electrochemistry*, P.E. White, B. E. Conway, and J. O'M. Bockris, eds., No. 30, p. 318, Plenum, 1996).

on the electrode. The two stages which this figure implies were elucidated largely by Kuzawa et al. in 1965.

In stage A:

$$MnO_2 + H_2O + e \rightarrow MnOOH + OH^-$$

In stage B:

$$MnOOH + H_2O + e \rightarrow Mn(OH)_2 + OH^-$$
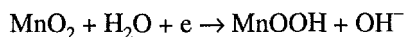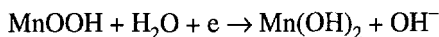
Advances were still being made in the 1990s concerning the mechanism of these reactions. In the first reaction (Stage A), the mechanism has been called the "proton-electron" mechanism and is portrayed in Fig. 13.42. The oxide is a mixture of Mn^{4+}, Mn^{3+}, O^{2-}, and OH^-, and the proton hops along between O^{2-} and OH^-, whereas the electron leaps onto Mn^{4+} from Mn^{3+}. This picture is a simplification and more details have been elucidated in terms of the MnO_2 structure, but our scope will be limited to the above.

In Stage B, the second electron step, the constant potential observed at low current densities indicates that the process does not involve a redox reaction subject to a Nernst equation, for, if it did, the potential would change according to the ratio of the redox
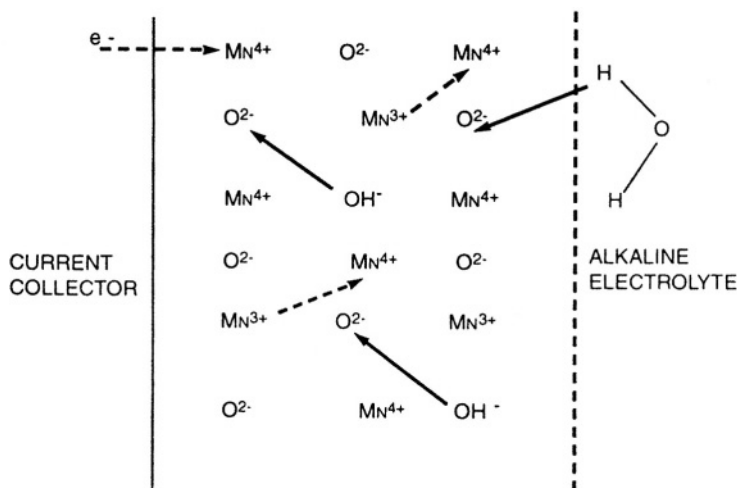
Fig. 13.42. Schematic representation of MnO_2 during the first electron discharge. (Reprinted from A. Kozawa and R. A., *Journal of Electrochemical Society* **113** (9): 871, 1966 with permission of The Electrochemical Society, Inc.)

species, which clearly changes as one gets converted to the other. It involves a dissolution-precipitation mechanism as follows:

$$MnOOH + H_2O \rightleftharpoons MnO(OH)_2^- + H^+$$

$$MnO(OH)_2^- + e_0^- + OH^- \overset{rds}{\rightarrow} Mn(OH)_4^{2-}$$

$$Mn(OH)_4^{2-} \rightleftharpoons Mn(OH)_2 + 2OH^-$$
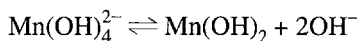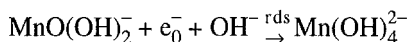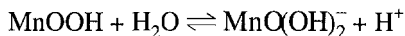
It will be seen that the cell's operation is far from simple.

Could $Zn-MnO_2$ cells become secondaries? In spite of the huge use of MnO_2 cells in flashlights, tape recorders, portable TV sets, etc., the cells have been found to be nonrechargeable (i.e., to be "primaries," thrown away after one use). However, they are low cost, have good energy and power densities, long shelf life and, above all, considering the present concern with environmental damage, have benign environmental properties. Hence, it has for long been realized that a rechargeable version of the $Zn-MnO_2$ battery would be a considerable plus in battery technology. It could replace Ni-Cd and even pose an alternative to the lead-acid battery.

A considerable clarification of the problem of rechargeability was made when it was realized that the refusal to accept several of the reaction sequences was connected with events in the second stage of the reaction. In starting out with MnO_2, and ending up with $Mn(OH)_2$, the oxides must pass through an average position of $MnO_{1.5}$. Now, when Mn^{4+} passes to Mn^{3+}, expansion of the lattice occurs. This expansion squeezes out the electrolyte from the interstices of the lattice, and when the current is reversed and the reactions of the cell are forced to go up the free energy gradient against their natural tendencies (i.e., in recharge), the lattice will not comply, for it no longer contains the necessary solution.

A partial solution to this problem was achieved by Kordesch. Why, he thought, provoke trouble? If this squeezing out effect does not occur until stage B in the full discharge of the battery, it would be wise to stop the reduction of MnO_2 at stage A. The one-electron reduction of MnO_2 would then have taken place. There would be no passage beyond $MnO_{1.5}$ and the trouble concerned with lattice expansion would not occur. Rechargeability of the one-electron reduction should then be possible. Realizing the virtues of the positive approach, half a loaf being better than no bread, Kordesch set about devising an MnO_2 cell which when discharged would stop at the half-way stage, at the end of stage A. He did this by two means. In the first, a device was contrived that could sense the characteristic potential (–0.5 V vs. an Hg-HgO reference electrode) at which stage B begins, and switch off the current. A simpler method was found to consist in starving the cell's Zn, so that by the end of stage A, all the Zn would be converted to ZnO and hence there would be no free energy left to continue on to

the second stage and reduce MnO_2 further than $MnO_{1.5}$. Such cells can be recharged to the extent of the one electron.

Although the Kordesch[30] philosophy is a good one, it is always possible to go one better. This was the aim of Halina Wroblowa,[31] a former member of the electrochemistry group that gave rise to Rex Watson, the electrochemist who produced the first practical fuel cell under the direction of Tom Bacon (Section 13.2) in the 1950s. Working in the Ford scientific laboratory with several collaborators, she concentrated her attention and that of her colleagues on the lattice disintegration caused by the introduction of Mn^{3+} (from Mn^{4+}) and OH^- (from O^{2-}). She also considered the possible reduction of semiconductivity in the second stage as an alternative or additional cause of the refusal of Zn batteries to accept recharge if discharged past 1e. Wroblowa et al. found birnesites and burserites to be the key to 2e rechargeability. These are allotropic forms of MnO_2 that have infinite 2D sheets of $Mn^{IV}O_6$ separated by 7–10 Å. It was argued in the Ford group that if they could introduce foreign cations into the birnesites and burserites, the open structure might be retainable, even after stage A had given rise to stage B, with its threatened lattice contraction and the electrolyte squeezing-out effect. This constructive thinking led to the discovery that, indeed, the introduction of Bi in the first place and Pb in the second, made a major change in the rechargeability of MnO_2 (Fig. 13.43).

Figure 13.44 shows that the high reactant utilization is maintained even down to 0.01 for the mole fraction of Bi_2O_3 in the MnO_2. The small amount of Bi_2O_3 that is effective in making possible 80% utility of the second stage in recharging MnO_2 (for up to 2000 cycles of deep discharge!) suggests that it may be the doping effect of Bi or Pb that allows a recharging of the cell after it has been discharged to the second

---

[30]In modern times there would be little controversy as to whom to name as the outstanding leader in battery research. It is Karl Kordesch, for long professor in an Austrian university but since 1986, director and the main force behind a Canadian company producing batteries. Among Kordesch's more visible contributions is his hybrid car of 1962 (a first for its time), which was part fuel cell (carrying cylinders containing hydrogen on the roof) and part battery driven. Several of the early analyses of the relation between fuel cells and batteries were published in papers with Kordesch as the leading author. A well-known analysis of the ultimate energy per unit weight with batteries is also due to him. However, his most lasting achievement is the so-called "rechargeable" MnO_2 cell described above. Ebullient and positive in attitude, Kordesch has made contributions to electrochemical storage that have been repeatedly fertile in concept and sustained over more than 30 years.

[31]Less well known to the battery community than Kordesch, Halina Stefany Wroblowa, originally from the University of Warsaw, Poland, came to prominence because she made Kordesch's 1e rechargeable MnO_2 electrode 2e rechargeable, an achievement that, if coupled successfully with Zn in a battery, could lead to a doubling of the capacity of the rechargeable cell. In classic research at Ford, Wroblowa identified two metals, which, when added to MnO_2, prevented the changes in the lattice that interrupted a successful 2e recharge (Section 13.16.2.3). Wroblowa, whose creativity is coupled to an unnerving ability to discover the frailties in others' hypotheses, was the first woman to chair the elite Gordon Conference in Electrochemistry. Her other firsts (from work at the University of Pennsylvania) include the discovery of the mechanism of electrochemical olefin oxidation, place exchange in the growth of passive oxides, and an accepted solution to the mechanism of the potential of zero charge on Pt in an O_2-containing solution.
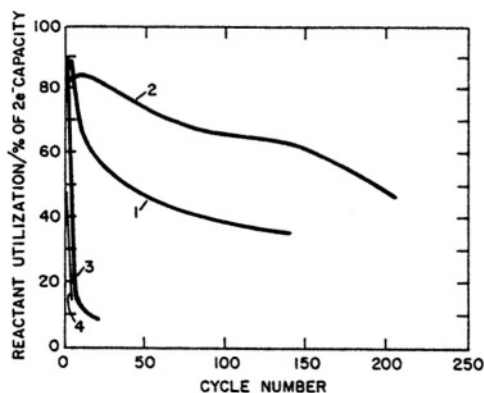
Fig. 13.43. Reactant utilization (percent theoretical two-electron capacity) vs. cycle number for 1, Pb-birnessite; 2, Bi-birnessite; 3, International Common Sample (ICS) no. 1; and 4, ICS no. 2. (Reprinted from Y. R. Yao, N. Gupta, and H. S. Wroblowa, *J. Electrochem. Soc.* **223**: 107, 1987. Reproduced with permission of The Electrochemical Society, Inc.)

stage. Thus, in addition to the squeezing-out effect of the electrolyte, it may be that material formed in the second stage of discharge is not sufficiently *electronically* conducting to allow passage of current throughout the potentially reacting mass of material on the electrode. *Doping* in semiconductors is effective at very low concentrations of dopant (compare the low concentrations of Bi_2O_3 that bring about the desired effect in recharging $Zn-MnO_2$). The original hypothesis that a lattice collapse "squeezes out" solution would seem to demand a far larger Bi_2O_3 to change the contracting power on the lattice upon the introduction of Mn^{3+}.

Now the 2e rechargeability of MnO_2 achieved by Wroblowa et al. does not yet mean that a rechargeable two-electron $Zn-MnO_2$ cell (the full loaf) has been obtained. When the $Bi-MnO_2$ is coupled with Zn, the 2000 cycles of rechargeability found when it is coupled with some inert counter-electrode is reduced to an insufficiently practical 50, allegedly because a Zn species diffuses through the membrane between Zn and MnO_2 on recharge and goes over to combine with the MnO_2, causing it to form zinc manganate. Improvements in the membrane (Conway, 1993) help, and an acceptable number of cycles may be on the way. Another solution may be to couple Wroblowa's 2e rechargeable MnO_2, not with Zn, but with Li in a rechargeable cell using nonaqueous solutions. Then, with no wandering zincate penetrating the separator and reducing the integrity of the MnO_2, Wroblowa's 2d producing MnO_2 may become a practical reality.
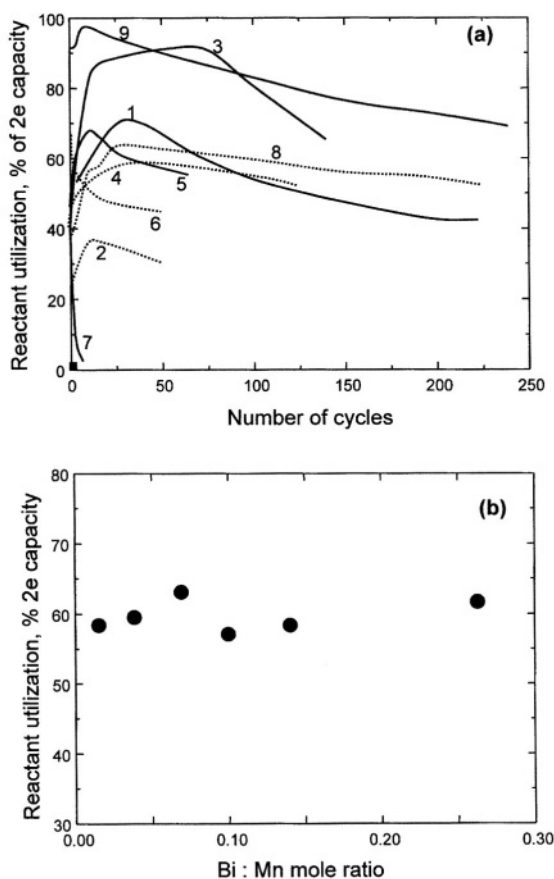
Fig. 13.44. (a) Reactant utilization as a function of the cycle number. 1 and 3, ICS no. 1 + Bi_2O_3; 2, ICS no. 1 + PbO; 4, ICS no. 4 + Bi_2O_3; 5, ICS no.7 + Bi_2O_3; 6, ICS no. 8 + Bi_2O_3; 7, ICS no. 1; 8, ICS no. 1 + PbO + Bi_2O_3; 9, Australian vernadite ore + Bi_2O_3. (b) Reactant utilization in the sixtieth cycle as a function of Bi-Mn mole ratio. (Reprinted from H. S. Wroblowa and N. Gupta, *J. Electrochem. Soc.* **238**: 93, 1997. Reproduced with permission of The Electrochemical Society, Inc.)

13.17.3. Modern Batteries

13.17.3.1. *Zinc-Air.* Zn-MnO_2 was discussed in Section 13.16.2.3. As a nonrechargeable battery, in various forms, its use is very widespread. As a rechargeable battery, it has its limitations at present. Zinc itself is an attractive battery material, being somewhat less corrodible in aqueous solution than aluminum. Thus, the relevant equation for the charging reaction with Zn is

$$ZnO + H_2O + 2e \rightarrow Zn + 2OH^-$$
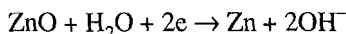
and the corresponding standard potential is –1.26 V. The question is, with what could one couple Zn to make a rechargeable battery that avoids uncertainties in the rechargeability of MnO_2? The cathodic dissolution of oxygen to form a partner to the anodic dissolution of Zn is a possibility. There are positive and negative aspects to this solution. On the positive side is the fact that using oxygen in air as one of the constituents of the battery reduces the battery's weight. Thus, one of the disadvantages with batteries, particularly where the watt hours per kilogram are so important, as in automotive applications, is that the active material is carried on the plates and the higher the capacity of a given battery, the greater the weight tends to be. If one electrode can be free of active material, i.e., if that material can be obtained from the surroundings, weight is saved. On the other hand, the exchange current density for the cathodic reduction of O_2 is particularly low (at best, 10^{-8} A cm^{-2} at 25 °C in alkaline solution) and hence there are likely to be substantial overpotential losses (see Section 13.8) and a lower watts kg^{-1} than with batteries involving reactions with higher exchange current densities.

Nevertheless, the Zn-air rechargeable battery is one of three batteries that are at the forefront of modern development. It is a (so-called) *secondary* battery, i.e., electrically rechargeable. The specific energy per unit weight is about 100 W hr/kg, or as much as 3 times better than the classical lead-acid cell, or twice that of the Ni-Cd cell. A detailed diagram of a practical version of this kind of battery is shown in Fig. 13.45. It is very important that the air electrode has two functions in such a cell. It has to act as cathode during the battery discharge cycle but as an anode during the charging. An $La_{0.6}$ catalyst [$Ca_{0.4}CoO_3$ (Ottagawa, 1981)] has been developed to optimize the evolution of O_2 in alkaline solutions. The oxygen (air) electrodes are double the apparent area of the zinc electrode and this halves the current density on them and lowers the overpotential. The bifunctional O_2 electrode coupled with Zn allows a fully rechargeable Zn-air battery; a discharge curve for a cell is shown in Fig. 13.46.

Another approach to the rechargeability of the Zn-air battery is to make the recharging mechanical. When the cell has discharged to (say) 80% of the maximal theoretical amount, the "zinc" electrodes (now largely ZnO) are removed and fresh ones inserted (e.g., in an automotive application).[32] The advantage is that the O_2 electrode now functions only as a cathode and does not have to be bifunctional. This

---

[32] Such a device was successfully demonstrated in a test vehicle by a U.S. automotive company in 1971.
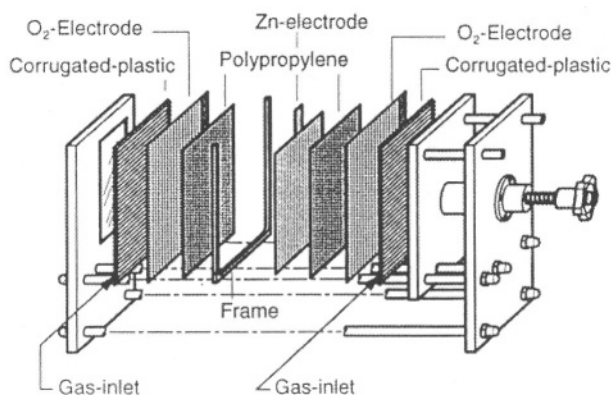
Fig. 13.45. Schematic view of the Zn-air battery consisting of a single cell with a central Zn anode facing two bifunctional air electrodes. (Reprinted from K. Müller, R. Holze, and O. Haas, "Progress Towards a 20 Ah/12V Electrically Rechargeable Zinc/Air Battery," in *Batteries for Portable Applications and Electric Vehicles,* C. F. Holmes and A. R. Landgrebe, eds., Electrochemical Society Proc. PV 97-18, pp. 859–868, Fig. 2, 1997. Reproduced by permission of The Electrochemical Society, Inc.)

lessens its cost and it turns out that the mechanically recharged battery has watt hr kg^{-1} of 200, double that with electrical recharging. A further advantage is that questions of the number of cycles in which the battery can be recharged (using a single piece of zinc) are no longer relevant. In the mechanically rechargeable version, the ZnO resulting from discharge has to be recycled back to Zn. If this is done at an electrochemical Zn recovery plant, the reaction can be achieved with a better efficiency than in the confines of a battery. Long impractical recharging times while the car is stationary are avoided.[33]

This battery has reached a commercial stage in Europe (1998) and offers a 200-mile range to electric cars. Of course, in countries where the approach is to be employed, Zn recovery plants have to be built. The mechanically rechargeable Zn-air battery may be compared with the mechanically rechargeable Al cell (Section 13.16.4.1). As will be seen, the latter offers a substantially larger energy density (watt hours kg^{-1} of 400–500) than the former. Just as ZnO has to be returned to the processing plant for electrochemical reduction to Zn, the Al(OH)$_3$, the product of discharge with Al, would have to be returned to a processing plant for electrochemical reduction to Al (an advantage is that plants for extracting Al from Al$_2$O$_3$ are already

---

[33]A mechanically rechargeable battery can be seen as a "primary," for it is not electrically rechargeable as a unit. The metal fed into it becomes a fuel and the battery is then like a fuel cell.
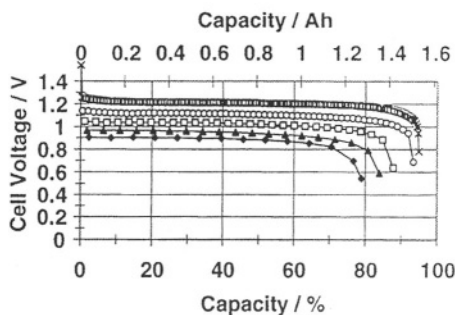
Fig. 13.46. Zinc–oxygen cell voltages and available capacity for different rates of discharge measured for two 25-cm^2 cells with pasted zinc electrodes containing 10 wt.% cellulose fibers with a fiber length of 1 mm. Discharge currents are ×, 0.064 A; ○, 0.3 A; □, 0.6 A; ▲, 0.9 A; ♦, 1.2 A. (Reprinted from K. Müller, R. Holze, and O. Haas, "Progress Towards a 20 Ah/12V Electrically Rechargeable Zinc/Air Battery" in *Batteries for Portable Applications and Electric Vehicles,* C. F. Holmes and A. R. Landgrebe, eds., Electrochemical Society Proc. PV 97-18, pp. 859–868, Fig. 3, 1997. Reproduced by permission of The Electrochemical Society, Inc.)

widespread throughout the world). Because of the lighter weight (Al, at. wt. 27; Zn, at. wt. 65), the mechanically rechargeable Al system offers a greater range than the mechanically rechargeable Zn-air cell.

13.17.3.2. *Nickel-Metal Hydride.*   According to the previous section, Zn as an anode and MnO_2 as a cathode in discharge are awkward partners for a secondary battery, but a mechanically rechargeable $Zn-O_2$ seemed to be feasible. Now, an analogous idea can be applied to nickel. For many years, the nickel-cadmium battery has been the leading rechargeable battery for toys and electronic devices. Cadmium, however, has been found to be toxic and the specific energy capacity of the Ni-Cd battery is only 1-1/2 times that of lead-acid. Thus, like Zn, nickel sought another partner. In batteries, on discharge, the partner that dissolves anodically is called "the negative" because its potential on the standard hydrogen scale (Sec. 7.2.7) is more negative than that of the partner electrode, which must thus act as the cathode (the positive), Now, Zn had chosen to couple with O_2. A contrarian alternative for nickel, therefore, would be a coupling with hydrogen. One would expect, now, that for

discharge, hydrogen would be the negative, dissolving anodically. However, the idea of coupling Ni with hydrogen, as an analogue of coupling zinc with oxygen, is not so easily implemented in a practical cell. Oxygen exists freely in the atmosphere and therefore does not have to be carried on the battery plates. Not so with hydrogen. Carrying it at high pressure in cylinders (which is worth considering if the cylinder material is a lightweight alloy of Al) is too heavy with the commonly used cylinders of steel. Nickel-hydrogen in an inconel pressure vessel was developed in 1970 for satellites, e.g., the Hubble telescope in low earth orbit and geostations for some telecommunications satellites. It gives 50,000 cycles at 40% depth of discharge. This is because the electrolyte (KOH) concentration remains constant during operation. This allows the choice of 28 wt.% KOH, which minimizes Ni corrosion. Starting about 1975, the idea of storing H in metals such as Ti or FeTi began to emerge. However, such systems are still too heavy. Nevertheless, one alloy was found which stored H to a far larger extent than did Ti, namely $LaNi_5$. However, this material had a poor cycle life. Then there occurred a discovery (Ovschinski, 1993). It was found that mixtures of certain metals (V, Ti, Zr) in alloy form have a splendid ability to absorb and store H when it is cathodically discharged onto them; they also have a good ability to let the H go again in a discharge cycle, in which the H dissolves anodically to become a proton in solution.

$$M_{alloy} H \rightarrow M_{alloy} + H^+ + e_0^-$$

So good were some of these alloys that their ability to absorb and redischarge H when coupled to some inert counter-cathode (not yet a battery!) was equivalent to $400 \, W \, hr \, kg^{-1}$. Thus was born the idea of the nickel-metal hydride battery, NiMH as it is written (where M is an alloy of V, Zr, Ti, etc.), one of the leading new batteries of the later 1990s.

What does the Ni side do in discharge? Here, the story is a bit complex and there is still much to be done with investigations of the mechanism. The essence of the cathodic reaction during discharge ("the positive") goes as follows:

$$NiOOH + H_2O + e- \rightarrow Ni(OH_2) + OH^-$$
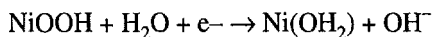
There is still much to learn about how to make such an electrode behave efficiently. It is not a simple matter. The $Ni(OH)_2$ (the form after discharge) is used as a slurry underlying metal plates. But thereafter, it must be properly sintered into small particle size and other materials (among them, cobalt) are added.

An advanced prototype cell shows a discharge curve as in Fig. 13.47. The value for the energy density of $94 \, W \, hr \, kg^{-1}$ given in the figure is in advance of the 60–70 generally measured for this battery. In this respect it gives a lesser specific energy density than the rechargeable Zn-air cell, but the 600 recharges for the cycle life and power density of $200 \, W \, kg^{-1}$ are encouraging and put the cell in the running for some automotive uses. The earlier automotive development of this battery seems to have been at Toyota, where as early as 1993 a car running on a test track was demonstrated for 300 miles at 75 mph on such a battery.
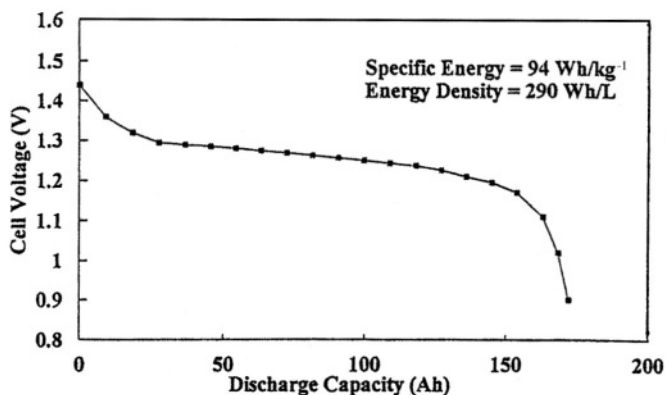
Fig. 13.47. Discharge curve for advanced prototype electric vehicle cell demonstrating 95 W h kg^{-1} cell-specific energy. (Reprinted from S. R. Ovshinsky, S. K. Dhar, S. Venkatesan, D. A. Corrigan, A. Holland, M. A. Fetcenko, and P. R. Gifford, "Ovonic Nickel Metal Hydride Technology for Consumer and Electric Vehicle Batteries-A Review and Update," in *Batteries for Portable Applications and Electric Vehicles,* C. F. Holmes and A. R. Landgrebe, eds., Electrochemical Society Proc. PV97-18, p. 706, Fig. 1, 1997. Reproduced by permission of The Electrochemical Society, Inc.)

13.17.3.3. *Li.*   Intrinsically, but naively, Li should have been regarded from the earliest days of battery science as the most attractive battery material because of its position as the most electronegative metal in the Electrochemical Series (Sec. 7.2.7). It has a thermodynamically reversible electrode potential of –3.0 V, and thus offers the greatest amount of electricity per unit weight among the elements. In the calculation of this quantity, H_2 is not compared directly with Li because of the need for a container, the weight of which would have to be taken into account. Dreams of the use of Li in batteries were at first entirely rejected because of its vigorous corrosive reaction in water, producing H_2, and the danger of explosions after contact of the H_2 with air. Li in organic solvents was also looked upon for many years as unrealistic because of the projected very low electrical conductivity of the electrolyte. Then, as usually happens in scientific developments, the imagined limitations were solved. Harris (1958) found that in many organic solvents (e.g., anhydrous butyrolactone, ethers) Li electrodes were perfectly stable; and fears concerning the poor conductance, while not groundless, had been exaggerated. During the 1970s, much progress was made in the use of Li as an anode material and by the 1980s, it was in use in a primary (i.e., nonrechargeable) cell for a variety of applications where a high energy density is needed. Li cells are generally sealed against the intrusion of air

because of the very strong interaction between Li and the water vapor it contains. The organic solvents used in these applications included ethers, propylene oxide, dimethyl formamide, propylene carbonate, etc. Many electrolytes have been used in these solvents and of course the all-important conductance will depend critically upon the solubility and ionicity of these compounds. They have tended to become relatively complex; examples are hexafluorophosphate and the corresponding arsenate.

In this brief account of Li cells, concentration will be more upon the moving frontier of rechargeable Li batteries which are being developed in many universities, industries, and government research laboratories, ostensibly with the possibility that one version might be viable as an automotive power source (although this aim has been blunted by developments in the late 1990s toward fuel cells for this purpose; see Section 13.7.1). Thus, the rechargeable cell works, obviously, on the basis of:

$$Li \rightarrow Li^+ + e$$

in the discharge cycle. The corresponding redeposition of Li onto a carbon matrix during cathodic charging occurs at 100% current efficiency, but the rate of the recharging reaction of Li^+ to Li has to be limited because if it approaches the limiting current density, dendrites form and there is the danger of membrane penetration, heating and short circuiting, and cell death.

The cathodic reaction during the discharge phase is more complicated, one of the difficulties of describing it being that it is being developed in several different versions. A description based on knowledge available in the late 1990s may be incomplete by 2005. At present, the cathodic reaction is carried out by reducing an oxide. Many oxides (V_2O_5, MnO_2, MOS_3) and sulfides (FeS_2)[34] are being used as cathodic materials. The cathodic reaction is in most systems rate-determining in respect to cell performance. The mode by which these oxides and sulfides are electrochemically reduced is complicated by the phenomenon of intercalation.[35] This means that the compounds themselves $(Fe_2O_3, Co_2O_2, MnO_2$, etc.) accept the Li^+ ions arriving from the anode (the anodic reaction being dissolution of Li), and then these ions diffuse $(Li^+$

---

[34]Also, certain organics, which seem to undergo both cathodic reduction and anodic oxidation, are being used.

[35]Intercalation is a word now used in a wider sense than before. It means "the insertion of something into another thing." In the Li batteries during discharge, the Li^+ produced at the anode diffuses to the cathode, which might be, say, TiS_2, and enters the lattice. Many of the systems to which the word "intercalation" was originally applied were layered structures, and the picture formed was that of Li^+ existing between the layers of the structure. However, the word is now broadened to conform to the above definition. There is some similarity between the way an Li-ion cell works (the Li^+ goes over to the cathode and enters it) and the working of the Ni-metal hydride batteries, in which H is intercalated inside the alloys of V, Ti, and Zr, etc., and ionizes out from it in discharge. Of course, the entry of Li^+ into the cathode and its diffusion thereafter is preliminary to the step of electron transfer *from* the cathode, and that is the part in which there is still some discussion (see above). Examples of formulas given for intercalated cathodes are $Li_xMoO_2; Li_xTiS_2, Li_xTiS_2, Li_xNb Se_2$, etc.

ions being of small size) within their interstices (the oxides are dispersed on and mixed with conducting carbon powders of various particle sizes and types). The real question, of course, is, what happens in the interstices of the oxides which accept the migrating cation (attracted to the cathode), for, finally, there must be a net electrochemical reaction from the conducting substrate to the oxide. It would be a gross misrepresentation, however, to think of this reduction of the oxide simplistically, e.g., as $CuO + 2H^+ + 2e \rightarrow Cu + H_2O$. To begin with, protons are not overtly present; and second, it is known that Li itself takes part in the reactions within the processes occurring inside the oxide-carbon mixtures that constitute the cathode. Represented in the presence of FeS_2, a probable cathodic sequence is (Strauss and Burstein, 1997)

$$Fe + Li_2S + 2Li^+ + 2e \rightarrow Li_2\,FeS_2$$

$$Li_2FeS_2 + x\,Li^+ + xe \rightarrow Li_{2+x}\,FeS_2$$
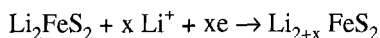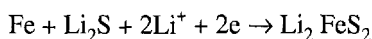
Alternatively (Oniciu, 1997), for a CuO-containing cathode,

$$Cu + xLi^+ + xe \rightarrow Li_x CuO$$

followed by

$$(2-x)\,Li^+ + (2-x)\,e_0^- + Li_2\,CuO \rightarrow Cu + Li_2O$$

$$(2-x)\,Li^+ + Li_x\,CuO + CuO + (2-x)e \rightarrow Li_2O + CuO$$

$$Cu_2O + 2Li^+ + 2e \rightarrow Li_2O + 2Cu$$
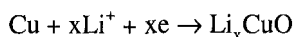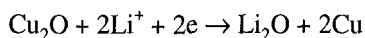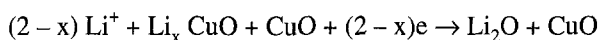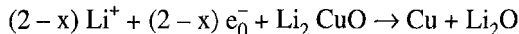
The reverse of these complex processes is assumed to occur on charging, but the investigations of the charging cycle have not yet reached a stable state.

Figure 13.48 shows a schematic of an Li cell in which the cathodic process involves an organic. Figure 13.49 shows how the capacity is affected by discharging and charging such cells more than 80 times.[36]

Finally, one must state some figures for the energy per unit weight of rechargeable Li cells, as well as their cyclability to an 80% depth of discharge. With so many varieties of cells described in the developing literature, only a range of values can be stated; this is 100–150 W hr kg^{-1}. The recyclability (number of possible recharges) will depend very much on the rates at which the cells are discharged. The cycle life is currently (1998) in the range of 400 for high to 1000 for low discharge rates.

A somewhat different type of rechargeable Li cell is being developed by workers who intend to use a solid polymer membrane as separator [see the use of solid polymer electrolyte (SPE) membranes in some fuel cells (Section 13.6.6)]. The advantage here

---

[36]The capacity expressed here is in terms of ampere hours kg^{-1}. If the cell had a constant potential of 2.0 V, the watt hours kg^{-1} would then be double the figures shown.
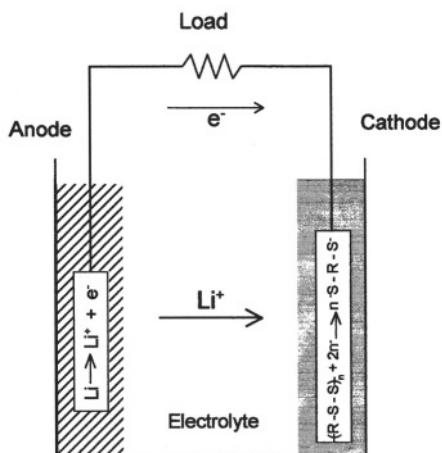
Fig. 13.48. Conceptual diagram of the discharge process of a lithium cell that includes an organosulfur-based cathode. (Reprinted from J. M. Pope, T. Sotomura, and N. Oyama, "Characterization and Performance of Organosulfur Cathodes for Secondary Lithium Cells: Composites of Organosulfur, Conducting Polymer, and Copper Ion," in *Batteries for Portable Applications and Electric Vehicles,* C. F. Holmes and A. R. Landgrebe, eds., Electrochemical Society Proc. PV97-18, pp. 116–123, Fig. 1, 1977. Reprinted by permission of The Electrochemical Society, Inc.)

is the absence of a bulk liquid. Polyethylene oxide has been suggested as the polymer membrane, and to be successful it must be permeable to Li^+ and *very* thin, or else the resistance will be too high and the watt hours kg^{-1} will become too low to be acceptable. An artist's diagram (not a schematic of a real cell) is shown in Fig. 13.50.

Finally, it must be stated that this account of rechargeable Li cells has been much influenced by the research literature up to 1998. The account should be qualified by the French proverb that says it is a great success to drive a bear up a tree, but you have to get it down again before you can claim the skin.

13.17.4. Some Batteries for Special Purposes

The Electrochemical Series of the elements (Sec. 6.3.13.3) is usually arranged in such a way that the most electronegative elements are at the top ($E_{o_1 Li}= -3.0$) and the
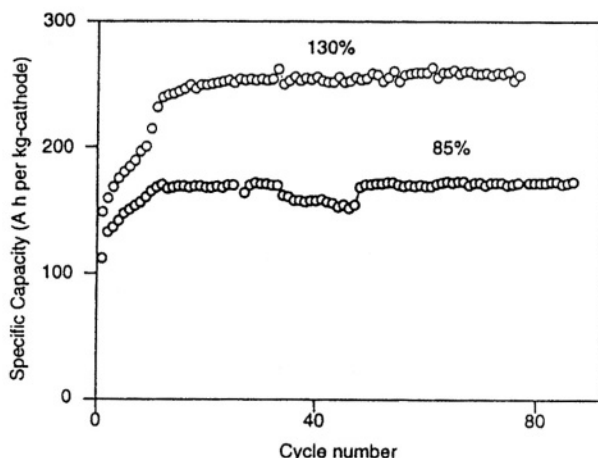
Fig. 13.49. Discharge capacity vs. cycle number for representative cells (at 20 °C) that were charged to 130 and 85% of their theoretical capacity. (Reprinted from J. M. Pope, T. Sotomura, and N. Oyama, "Characterization and Performance of Organosulfur Cathodes for Secondary Lithium Cells: Composites of Organosulfur, Conducting Polymer, and Copper Ion," in *Batteries for Portable Applications and Electric Vehicles,* C. F. Holmes and A. R. Landgrebe, eds., Electrochemical Society Proc. 97-18, p. 122, 1997. Reproduced by permission of The Electrochemical Society.)

most electropositive are at the bottom ($E_{O_1F_2} = +2.86$). Without much thought, one might assume that an Li-F_2 cell would have the greatest possible cell potential and perhaps also storage capacity because Li is light (see Table 13.9). However, as indicated, Li reacts violently with water and F_2 is also not unreactive. Going down the list from the most electronegative toward H_2 (in the middle, $E_{O_1H_2} = 0$), one comes across aluminum at −1.66 V.

What about aluminum as a store for electrical energy? It is relatively light ($\rho = 2.7$) and its dissolution gives rise to 3 electrons per atom so that per electron it is comparable with Li in electrical energy per unit weight (Table 13.9). On the other hand, although it corrodes in aqueous alkaline solution when not connected up with a partner electrode to form a cell, the corrosion is relatively slow, and can be inhibited by adding the appropriate organics to the solution (Section 12.3.2). Aluminum is abundant in nature in the form of clay (sodium aluminum silicate) and bauxite (hydrated alumina). Solomon Zaromb (1960) first proposed using available electricity to extract Al from its ores, store the Al in the form of thin sheets,
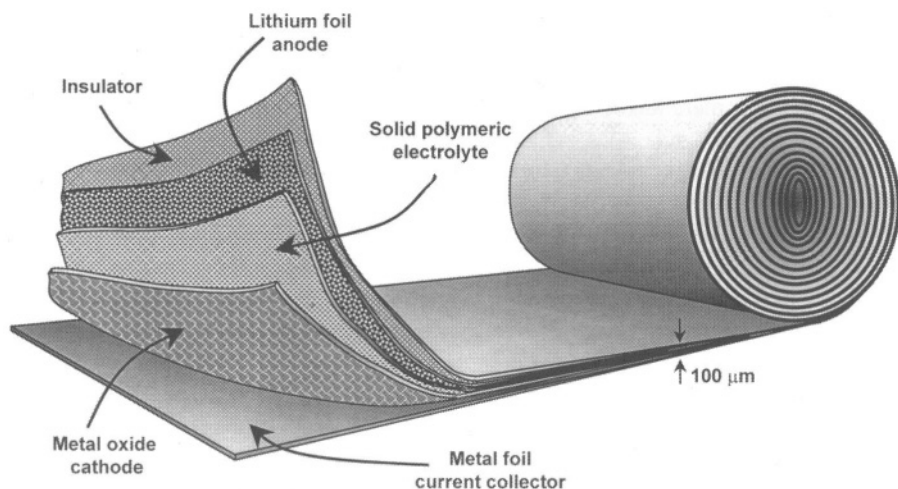
Fig. 13.50. A solid-state Li polymer battery in conception. (Reprinted with permission from S. L. Wilkinson, *Chem. & Eng. News*, Oct. 13, 1997, p. 23).

and re-obtain the electricity upon demand from batteries consisting of Al coupled to an O_2 electrode.

During discharge:

$$Al \rightarrow Al^{3+} + 3e \qquad \text{(negative)}$$

$$\frac{\dfrac{3}{4} O_2 + \dfrac{3}{2} H_2O + 3e \rightarrow 3OH^-}{Al + \dfrac{3}{4} O_2 + \dfrac{3}{2} H_2O \rightarrow Al^{3+} + 3OH^-} \qquad \text{(positive)}$$
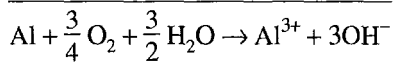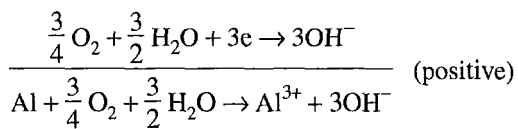
TABLE 13.9
Electricity Storage per Unit Weight of Active Metals

Metal	Electric Charge per Gain on Anodic Dissolution
Li	0.14
Be	0.22
Mg	0.08
Zn	0.03
Al	0.11

The Al^{3+} in alkaline solution then forms a white fluffy-looking precipitate of $Al(OH)_3$, which can be arranged to descend into a suitable space below the electrodes, the material being recycled in an aluminum extraction plant. "Aluminum batteries" had not yet been commercialized by 1998, but they stimulate interest in two directions.

1. They may provide a safe, light storage material for automotive transportation. Just as there are safety advantages in storing energy in methanol[37] rather than in hydrogen for fuel cell-powered transportation (Section 13.7.1), storage of energy in aluminum represents the ultimate in safety in respect to danger from explosion or fire in automotive transportation. The weight is further lessened because there is no need for a container. Further, it is easy to design methods of carrying Al (within the volume of a present gas tank) in the form of rolls of thin foil which, when fed into an Al-air cell, would drive a car at 25 kW for more than 1000 miles.

2. They could serve as reserve batteries to be used with an electrolyte of sea water. Despic (1981) utilized the properties of Al alloys with tin to shift the potential of Al in the negative direction and increase the potential of the cell formed with O_2. The Cl^- ion in sea water breaks down protective layers that would reduce the rate of anodic dissolution in an Al-O_2 battery. In fact, in sea water, such batteries can function at the extraordinarily high rate of 1 A cm^{-2}. The watt hours per kilogram at low rates of discharge are 500, which is well above the practical range for other batteries. In view of the commercialization of mechanically rechargeable Zn-air batteries for automotive applications, the commercialization of Al batteries in the United States is conspicuous for its slowness.[38]

13.18. THE VIEW AHEAD WITH BATTERIES

13.18.1. General

There is only one way in which electricity can be stored and then instantly obtained by closing a switch. In this respect, the electrochemical electricity storer, a battery, has no rival, and fuel cells in particular do not threaten batteries, for they have a different purpose—to create electricity from chemicals held in receptacles near the cell.

In spite of the great acceleration in battery development since the 1970s, there is still a large gap between the energy storage density readily available (about 100 W hr kg^{-1}) and the theoretical maxima. The latter (see Fig. 13.51) reaches about 500 W hr kg^{-1} for cells using aqueous solutions at room temperature, and 2000 W hr kg^{-1} for the high-temperature (molten salt) batteries (Fig. 13.52).

---

[37]However, methanol is toxic if ingested.

[38]The cell has a downside that it is economically attractive only if the $Al(OH)_3$, the product of the discharge reaction, can be recycled. A similar fact applies to the mechanically rechargeable Zn cells. The cell is not electrically rechargeable (though Al can be plated out from a nonaqueous solution or easily recovered by the normal method of the electrolysis of an Al_2O_3 cryolite mixture). However, mechanical recharging can be carried out at a faster rate and less expensively.
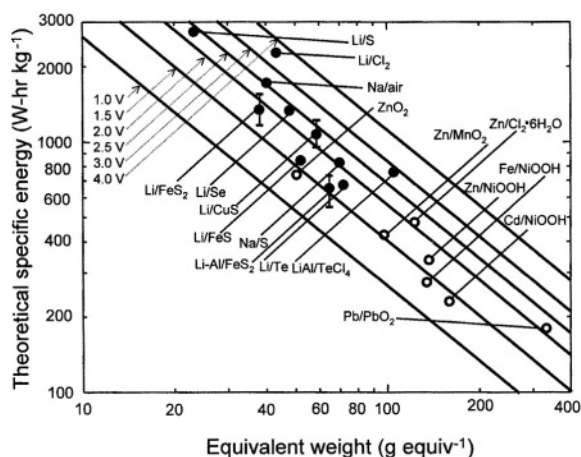
Fig. 13.51. Theoretical specific energy plotted against the equivalent weight for various batteries. The present commercial battery systems are in the lower right corner. Types of electrolytes: ●, molten salt or ceramic; ○, aqueous. (Reprinted from K. Kordesch, in *Comprehensive Treatise of Electrochemistry,* J. O'M. Bockris, B. E. Conway, E. Yeager, and R. E. White, eds., Vol. 3, p. 123, Plenum, 1981.)

Figure 13.51 shows how poor the performance is of the batteries already commercialized (~ 50 W hr kg^{-1}) and emphasizes systems that are open for research and development. Figure 13.52 shows similar information but as a function of time. The proliferation of new types of batteries since 1970 is noteworthy. However, some of the newly engineered batteries do not yet have sufficient cycle life to be widely usable.

The superior watt hours kg^{-1} offered by fuel cells (see Fig. 13.52) must again not be taken as a threat to batteries. The realm of effectiveness of batteries encompasses situations where it would be impractical to store fuel to make electricity on the spot. Such situations are widespread, particularly in powering portable equipment, from telephones to tape recorders.

13.19. ELECTROCHEMICAL CAPACITORS AS ENERGY STORERS

13.19.1. Introduction

In the parts of this chapter dealing with storage of electrical energy, it has been shown that electrochemical electricity storers, batteries, have certain properties (see Table 13.8), the most important of which is in the energy density (30–300 W hr kg^{-1}).
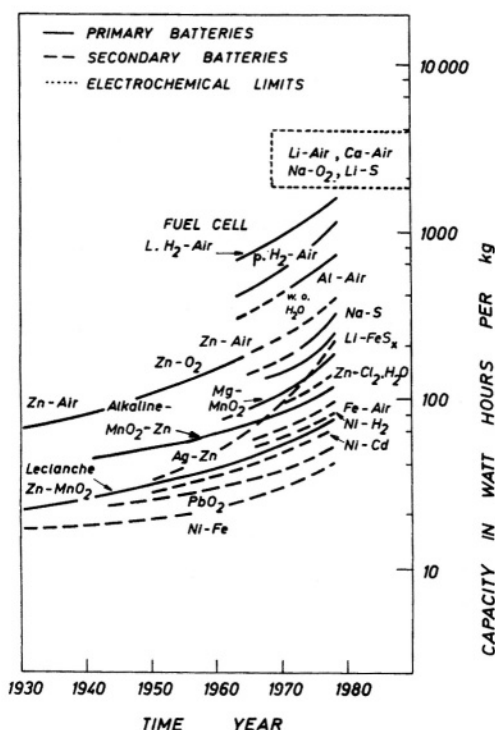
Fig. 13.52. Improvement of battery systems be-
tween 1930 and 1977 and the capabilities of
advanced systems. (Reprinted from K. Kordesch,
in *Comprehensive Treatise of Electrochemistry*,
J. O'M. Bockris, B. E. Conway, E. Yeager, and R.
E. White, eds., Vol. 3, p. 123, Plenum, 1981.)

It has long been realized that there is an altogether different method of storing
electrical energy. It is called *dielectric storage* and consists, esentially, of building
up electrical charges on two metal sheets separated by a dielectric. Such devices
can be recharged very rapidly and the potential across the dielectric can amount to
$\sim 10^3$ V. However, when one calculates the watt hours kg^{-1} available from such
devices, it is miserably small, less than 0.1 W hr kg^{-1}, which is no competition for
the electrochemical battery, though it is of great use if a large pulse of electricity
lasting a very short time is needed. The most simple equation for the capacity of
a parallel-plate condenser (Sec. 6.6.2) is

$$C = \frac{\varepsilon}{4\pi\delta}$$

per unit area, where ε is the dielectric constant and δ the distance between the plates. It is obvious that to increase the capacity, and hence the energy storage density, one needs to make δ as small as possible. This is achieved in practical condensers by using thin oxide films (Ta_2O_5, for example). One may conceive that δ could be as little as, say, 0.1 μm or 10^{-5} cm.

Is there a way of making δ much less even than this? The answer is yes, and one is led directly to *electrochemical* capacitors. The interfacial region at the metal/solution interface has been well studied (Chapter 6) and it is known that the value of δ for the capacitance of the double layer is on the order of a molecular diameter, say, 5 Å or 5 10^{-8} cm. Thus, the capacitance of the polarizable interface is $10^{-5}/5 \times 10^{-8}$ or 200 times greater per unit area than that of a dielectric capacitor. The value of the dielectric constant would be greater for an oxide HO (~20) than for the interfacial region (~6), but this still leaves a ratio of ~10 times in favor of normal batteries. However, although the watt hours kg^{-1} is the primary charac- teristic by which one classifies batteries, it is by no means the only one (Section 13.15.3). The *power* density of electrochemical capacitors is on the order of 10^2 kW kg^{-1}, which is 100–1000 times more than that of batteries! The cycle life is about 10^5, again immensely better. This makes one pause and take heed. In insisting on batteries with their electrochemical reactions and the complexities that these bring, are we putting our funds on the wrong projects? Before discussing this, let Fig. 13.53 summarize what has been said so far.
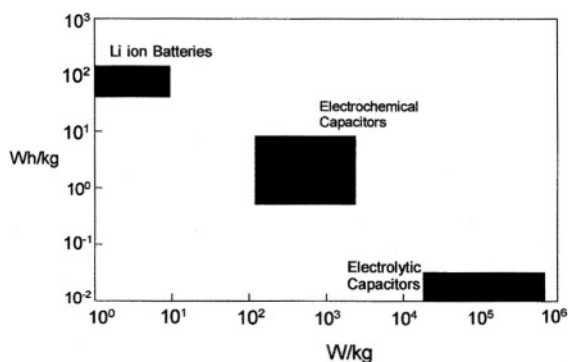
Fig. 13.53. Ragone plot comparing electrochemical capacitors with lithium ion batteries and electrolytic capacitors. (Reprinted from Electrochemical Society Proc. 97-18, p. 610, Fig. 1, 1997. Reproduced by permission of The Electrochemical Society Inc.)

13.19.2.   Can the Energy Storage Possibilities with Electrochemical Condensers be Greatly Increased?

The properties of electrochemical capacitors summarized in the previous section are not the only ones that stimulate further discussion of these devices. Batteries work by means of a series of electrode reactions, nearly always involving electrochemical crystallization. The difficulties this causes came to the fore in the discussion of the MnO_2 cell in respect to its difficulties on recharge, and they also lurk in the complexity of the cathodic reaction on discharge of the Li-ion cells. There would be nothing of this kind in electrochemical capacitors, for there only electrons move in and out of the metal and ions and rearrange themselves in the interfacial region. Further, since there is no chemical reaction, there is no entropy change and the resulting heating or cooling.[39]

There are three routes by which the capacitative energy storage can be improved:

1. Achieve a highly porous structure, resulting in a much greater real area per square centimeter. What one needs is a compressed powder so that in fact the apparent unit area turns out to be 100 or even $1000 \, cm^2$ internally per apparent external square meter. Some idea of what is needed on the porosity side can be seen from Fig. 13.54.

2. The capacity per unit area that arises at a purely polarizable electrochemical interface (using the equation $C = \varepsilon/4\pi\delta$), is about $20 \, \mu F \, cm^{-1}$. The energy storage of such an electrochemical capacitor is proportional to this value. However, one may be able to involve the phenomenon of pseudo-capacitance to increase the capacity above that for a polarizable interface. Then one could envisage an electrochemical ultracapacitor with capacities of $150–200 \, \mu F \, cm^{-2}$, an increase of up to 10 times that available from the capacity of polarizable interfaces.

3. An increase in energy densities is feasible through the use of nonaqueous systems, particularly the pure liquid electrolytes at room temperature (Vol. 1, Section 5.12), such as certain tetraalkylammonium nitrates. The advantage (McEwan, 1997) is that it greatly widens the range of potentials (3–4 V) over which the electrochemical capacitor can be ranged without running into the danger of evolving H_2 on the negative side or O_2 on the positive.

Certain noble metal oxides (e.g., RuO_2 and Ir_2O_3) are particularly suited for electrochemical ultracapacitors. They can be grown electrochemically in a very high area ("black") form or made chemically as powder. Of course, the help that porosity is assumed to give is only realizable if the oxides have sufficient electronic conductivity, and that requirement is fulfilled for the oxides named. As to pseudo-capacitance, that involves some charge exchange ("leaky condenser") and can be seen to be a possibility with RuO_2, which undergoes reactions of the following type (Trasatti, 1991) when subject to potential change in contact with aqueous solutions:

$$RuO_x(OH)_y + \delta \, e_0 + 8H^+ \rightarrow RuO_{x-8}(OH)_{x+\delta}$$
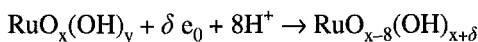
[39]In fact there is some, but it is ~1% of that for batteries.
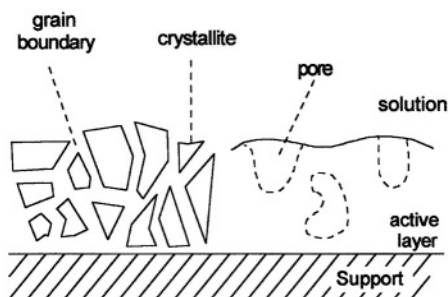
Fig. 13.54. Morphology of a thick oxide layer, showing crystallites, grain boundaries, and pores. The oxide/solution interface area is very large when the layer is impregnated with electrolyte solution. (Reprinted from S. Trasatti and K. Weil, "Electrochemical Supercapacitors as Versatile Energy Stores," *Platinum Metals Rev.*, **38** (2): 53, Fig. 8, 1994, with permission from Johnson, Matthey & Co.)

It can be seen from the way this reaction is written that the charge exchange is concerned with hydrous structures (Burke, 1990).

13.19.3. Projected Uses of Electrochemical Capacitors

Two main types of use are foreseen for electrochemical capacitors. They will be preferable to batteries in any application in which high power density, fast response, rapid charging, and high cycle life are more important than energy density. One might think of cellular telephones and memory protection systems in computers. Very little energy is needed at one time, but it has to be switched on and off very rapidly, be utterly reliable, and not need renewal for many years. The other possibility is as a booster power source for other power sources (e.g., Li-ion or fuel cells) where a ready and reliable supply is needed, the capacitor coming into play for a 1-minute boost of special power needed at startup and overtaking. Emergency electromagnetic braking might well be a use, particularly on very large trucks (18 wheelers), where the need for a new reserve braking capacity is felt by many road users. Because the intensive development of these devices is a matter of the later 1990s, it is not yet time to make a final judgment on the utility of electrochemical capacitors. The vital question, of course, is whether they could ever compete with batteries on the energy density front. Figure 13.53 shows that one would have to increase the storage capacity of the condensers by about 100 times to make it possible to replace the one electrochemical device with the other. In principle, an increase in this amount in terms of higher

porosity is possible. Here, again, one needs to investigate the conductivity of surface oxides (e.g., of RuO_2). It is mainly the conductance of the solid that counts, for no electrochemical current is to pass through the pores, so that the electrolyte conductance in them is of lesser consequence.

13.20. BATTERIES: AN OVERVIEW

When looking at the battery scene in the United States, Japan, and Europe around the end of the twentieth century, one sees a vibrant, extremely active, and developing industry that is fully alive on the research side and that has several frontiers on which it is growing. Much of this is connected with the environmental aspects of life. Because of their application in so many different directions, there is no "best battery," but there are batteries of different types for different situations. One has to have batteries for heart pacers, but one also makes batteries to run torpedoes. There is only a superficial similarity between batteries and fuel cells. Their aims, of course, are entirely different. One stores electricity produced elsewhere and the other produces electricity on the spot. Occasionally, a battery becomes virtually a fuel cell. Thus, the mechanically rechargeable Zn-air battery, though formally a battery might be regarded as functioning like a fuel cell.

Batteries have been with us for a long time, and three of the classical batteries that are studied in this chapter (lead-acid, nickel cadmium, and zinc manganese dioxide) were all invented in the nineteenth century. The scheme in the chapter has been to present these classical types, but include advances that may make zinc manganese dioxide also rechargeable; then to look at three modern batteries: Zn-air, nickel-metal hydrides, and lithium-ion batteries. After that, a battery not yet commercialized (aluminum-air), was presented and finally, a new type of electrochemical storer, the electrochemical capacitor. There are several key properties by which one determines the capabilities of a battery. Its discharge plot is its potential vs. time relation and should be as flat as possible for the whole course of the discharge. The Ragone plot is the plot of specific power vs. the specific energy, and the energy that is usable declines as the rate of use increases. Thus, we cannot state that a given battery has an outright single specific energy and power.[40] Each depends upon the other.

Energy density is the most important characteristic of the battery, and the energy densities from available aqueous batteries vary widely, from a low of 30 W hr kg^{-1} to a high of about 300. But if one looks at molten salt batteries, perhaps part of the future, then we are looking at up to 1000 or even $2000 \text{ W hr kg}^{-1}$ as a possibility. Charge and discharge behaviors for batteries are remarkably different because overpotential takes a bit out of the available potential in both directions. Thus, when one applies a potential to charge a battery at a given rate, it has to be equal to the reversible potential plus the sum of the polarization voltages and IR drops. However, when one takes energy from

---

[40]Although such figures are given, using the value that corresponds to the middle ranges of normal use.

the battery, then the maximum that one can get (limitingly low rates) corresponds to the thermodynamic potential, and as one increases the rate of discharge to a practical range, the overpotential pulls the available energy down, below the amount thermodynamically available in the reversible zero current state.

The lead-acid battery has a high power density, but is among the poorest in energy density. It is under our nose, so to speak, because it has found an ideal niche in its use in starting cars. Nearly all batteries have a better, some of them very much better, energy per unit weight than the lead-acid battery. The nickel-cadmium cell is a little better than the lead-acid battery charge in energy density, but its main advantage is that it has a much better cycle life (although it presents an environmental hazard, toxic cadmium). The zinc-manganese dioxide battery is the battery that powers most of our solid-state devices. In its nonrechargeable form, it is the power source for applications from flashlight batteries to portable TVs and computers. Attempts to make the zinc manganese dioxide battery rechargeable make an interesting story that is not yet finished, although it now seems that certain forms of manganese dioxide can be fully (using two electrons per mole) recharged if they are not coupled with zinc but with some other counter-electrode, such as lithium.

Among the modern batteries is Zn-air, giving a reasonable 100 W hr kg^{-1}, and now being applied for automotive use in Europe in a semicommercial way. When it is mechanically recharged, it gives 200 W hr kg^{-1}. Much newer than this battery is the nickel-metal hydride battery that was developed during the 1980s and 1990s. This battery is based on the discovery that vanadium, titanium, and zirconium, in alloy form, store hydrogen much more efficiently than previously used systems. These alloys absorb hydrogen when it is cathodically evolved upon them in the charging reaction, and the hydrogen dissolves out of them anodically during discharge. Meanwhile, over at the nickel electrode during discharge, nickel hydroxide is being formed from $NiOOH$.

The frontier of battery research in the late 1990s was in the lithium cells, using nonaqueous solutions to avoid Li's instability in aqueous solutions. There are both nonrechargeable and rechargeable Li batteries. We are interested here mainly in the rechargeable ones because they offer possibilities as a power source in electric vehicles. They use a nonaqueous solution (for example, butyrolactone) and an electrolyte (for example, hexafluorophosphate). The answers to two questions for lithium ion cells remain very much in the laboratory. What happens at the cathode during discharge, and can solid polymer electrolytes be developed to make the cell free of bulk liquids? As far as the cathode is concerned, the principal phenomenon is intercalation. The lithium comes over from the dissolving anode during discharge and enters into the interstices of the oxides, such as CuO, which are dispersed upon some conducting substrate, such as carbon. The electron-transfer reaction that must occur happens complexly after a series of reactions that are not yet entirely agreed upon. If it were practical to have a solid polymer that is a lithium conductor, the possibilities

for this cell would be increased, just as the proton conducting membranes made the fuel cell viable as an automotive power source.

Finally, in our discussion of electrochemical batteries we focused on aluminum as a battery material. It is not rechargeable electrically and if it were used extensively, particularly in cars, it would have to be mechanically recharged, just as the zinc-air cell. The great advantage is that it is so light that the range it would give a car would be comparable with that promised by the BMW company using pure liquid hydrogen in a fuel cell.

The final part of our chapter turned to something quite new in the world of electricity storage, namely, the use of condensers for this purpose. When high-voltage pulses are needed for short-term discharge, this technology has long been used. However, when energy density is important, the normal dielectric condensers give hopelessly small values for this very important quantity. It is then that one recalls that the electrical double layer is itself a condenser, but with a distance of a few angstroms between "the plates." One realizes that the capacity of the electrochemical double layer has possibilities for electricity storage. Condenser storage offers tremendous power density, and the number of cycles of recharge is so great that it ceases to be a consideration. It is true that even using an electrochemical interface, such a device has an energy density that is too small for it to be compared with other devices, even with the lead-acid battery, but there are applications (for example, cellular telephones) where the energy density of the condensor storers is sufficient and their reliability, a long lifetime, and power density come to the fore. This chapter describes various properties of such devices (e.g., electrochemical ultracapacitors) which might, in the coming decades, make it possible for electrochemical condenser storage to challenge even lithium-ion power sources.

Further Reading

Seminal

1. Plantè: *Recherches sur l'electricite,* Paris (1879). The invention of the lead-acid battery.
2. K. Kordesch in *Comprehensive Treatise of Electrochemistry*, J. O'M. Bockris, B. E. Conway, E. Yeager, and R. E White, eds., Vol. 3, p. 123, Plenum, New York (1981).
3. *Encyclopaedia Brittanica,* Vol. 7, p. 231, 1987. Leclanchè invented the Leclanche battery in 1866.
4. F. Gassner, German Patent 37,748,1888. This added paste to NH_4Cl solution in the $ZnOMnO_2$ battery. Hence the phrase "dry cell."
5. M. Barak, in *Comprehensive Treatise of Electrochemistry*, J. O'M. Bockris, B. E. Conway, E. Yeager, and R. E. White, eds., Vol. 3, p. 191, Plenum, New York (1981). Primary cells.
6. M. Kronenberg and G. Blomgren, in *Comprehensive Treatise of Electrochemistry*, J. O'M. Bockris, B. E. Conway, E. Yeager, and R. E. White, eds., Vol. 3, p. 247, Plenum, New York (1981). Primary cells for special purposes.

7. E. Kordesch, in *Comprehensive Treatise of Electrochemistry,* J. O'M. Bockris, B. E. Conway, E. Yeager, and R. E. White, eds., Vol. 3, p. 127, Plenum, New York (1981). Rechargeable batteries.

8. H. Andrè, *Bull. Soc. Fr. Electr.* **1**(6): 132 (1941). Development of a Zn-Ag cell.

9. A. Kozewa, in *Batteries,* E. Kordesch, ed., Vol. 1, p. 385, Marcel Dekker, New York (1974). Mechanism of MnO_2 reduction.

Modern

1. Y. F. Yao, N. Gupta, and H. S. Wroblowa, *J. Electroanal Soc.* **223:** 107 (1987). Studies on the effects of adding Bi and Pb to the two-electron rechargeability of Zn-MnO_2 cells.

2. M. A. Dzieciuch, N. Gupta, and H. S. Wroblowa, *J. Electroanal. Chem.* **138:** 2416 (1988). Coupling of 2e rechargeable MnO_2 with Zn.

3. P. A. Fiedler and J. S. Besenhard, *Proc. Electrochem. Soc.* **18:** 893 (1977). Voltammetric characterization of MnO_2 as affected by metal ions.

4. M. Kloss, C. Gruhnwald, D. Rahmer, and W. Plieth, *Proc. Electrochem. Soc.* **18:** 905 (1997). On the rechargeability of MnO_2.

5. K. Kordesch and C. Feistaner, *Proc. Electrochem. Soc.* **97:** 923 (1997). Rechargeable alkaline MnO_2 cells in practice.

6. E. Strauss, D. Golodnitsky, Y. Lavi, E. Peled, L. Burstein, and L. Lateah, *Proc. Electrochem. Soc.* **18:** 133 (1997). Rechargeable Li-FeS_2 cells for electric vehicles.

7. J. Barker, J. Swoyer, and M. Y. Caidi, *Proc. Electrochem. Soc.* **18:** 241 (1997). Li-V_6O_{13} cells.

8. C. B. Appetechi, F. Croce, B. Scrosati, and M. Wahihera, *Proc. Electrochem. Soc.* **18:** 488 (1997). Li-MnO_2 cells separated by plastic ionic membrane.

9. J. Flores, M. Urguidi-MacDonald, and D. D. MacDonald, *Proc. Electrochem. Soc.* **18:** 528 (1997). Li coupled with Pt electrode makes a powerful cell in aqueous solutions.

10. S. R. Orshinski, S. K. Dhar, S. Venkatesen, D. A. Corrigan, A. Holland, M. A. Fetsenka, and P. R. Gifford, *Proc. Electrochem. Soc.* **18:** 693 (1997). A review of metal hydride batteries stressing U.S. work.

11. S. Gamburzev, O. A. Velev, R. Danin, S. Srinivasan, and A. J. Appleby, *Proc. Electrochem. Soc.* **18:** 726 (1997). Improved design of nickel-metal hydride batteries.

12. Y. Yang, J. M. Nan, J. Ki, and J. G. Lin, *Proc. Electrochem. Soc.* **18:** 750 (1997). Surface properties of metal hydrides.

13. A. B. McEwan, J. L. Goldman, T. Blakey, W. F. Averil, and V. R. Koch, *Proc. Electrochem. Soc.* **18:** 602 (1997). Nonaqueous electrochemical capacitors.

14. D. M. Wilde, T. J. Guther, R. Oesten, and J. Gorche, *Proc. Electrochem. Soc.* **18:** 613 (1997). Perovskites as the basis of electrochemical supercapacitors.

15. L. Redey, *Proc. Electrochem. Soc.* **18:** 623 (1997). Heat effects in electrochemical capacitors.

16. "Environmental Concerns, Public Policies, and Remediation Technologies," in *Proc. Third Ann. Conf. on Environmental Science,* R. J. Gale, W. J. Catallo, R. C. Mohanty, and J. B. Johnston, eds., *American Society for Environmental Science,* Baton Rouge, LA (1993).

17. F. M. Delnick, D. Ingersol, X. Andrieu, and K. Naoi, eds., *Electrochemical Capacitors, Proc. Electrochem. Soc.* **96–25.** *A* collection of papers.

EXERCISES

1. (a) Sketch out the plot of efficiency vs. current density normalized to show the limiting current for an idealized fuel cell (two flat plates in a cell). (b) Draw a working potential vs. current density plot for an idealized fuel cell and label the various parts of the diagram as to the dominant modes of energy loss in each part. (Bockris)

2. Consider two half-cell reactions, one for an anodic and the other for a cathodic reaction. The exchange current densities for the anodic and the cathodic reactions are 10^{-6} A/cm^2 and 10^{-2} A/cm^2, respectively, with transfer coefficients of 0.4 and 1, respectively. The equilibrium potential difference between the two reactions is 1.5 V. (a) Calculate the cell potential when the current density of 10^{-5} A/cm^2 flows through the self-driving cell, neglecting the concentration overpotentials. The solution resistance is 1000 Ω cm^2. (b) What is the cell potential when the current density is 10^{-4} A/cm^2? (Kim)

3. Fuel cells convert chemical energy to electricity with an efficiency that depends (largely) on the overpotential of the oxygen cathode. They are not affected by the Carnot efficiency limit, as are combustion engines. (a) Examine in terms of plots the variation of the Carnot efficiency expression with T_{high} for a number of reasonable T_{low} values. (b) Examine the fuel cell efficiency expression $(V_{rev} - \Sigma\eta)/\Delta H$ with variation of $\Sigma\eta$ and for an H_2-O_2 cell. Find out how high the T_{high} of a combustion engine would have to be for a fuel cell efficiency above 0.6. (b) What magnitude of $\Sigma\eta$ would reduce the fuel cell efficiency to below 0.5? (Bockris)

4. In the operation of a conventional PEM fuel cell, the dry gases (hydrogen-oxygen) have to be humidified before entering the fuel cell. Generally, they are passed through water bottles and humidified with water vapor. As soon as the humidified gases enter the fuel cell, the vapor will condense and release heat. You're running a fuel cell under 1 atm and 60 °C. The flow of gas is 100 liters/hr with saturated water vapor at 1 atm. (a) Calculate the heat released from the water vapor (kJ/hr). (Ho)

5. The molten carbonate fuel cells employ Li_2CO_3-K_2CO_3 (62.38 mol.%) electrolytes, porous Ni alloy, and lithiated nickel oxide as anodes and cathodes at an operating temperature of 723 K. The half-cell reactions of each side are, respectively

$$\text{Anodic reaction: } H_2 + CO_3^{2-} \rightarrow H_2O + CO_2 + 2e^- \quad (H_2 \text{ fuel})$$

$$CO + CO_3^{2-} \rightarrow 2CO_2 + 2e^- \quad (CO \text{ fuel})$$

Cathodic reaction: $1/2 O_2 \mid CO_2 + 2e^- > CO_3^{2-}$
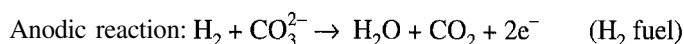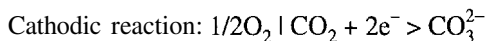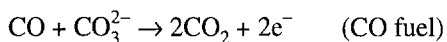
None of the oxidation reaction of CO is taken into account in this problem. The single-cell (electrode area 0.5 m^2) of 1-mm-thick electrolyte plate operates at a pressure of 7 atm. For the anode fuel gas the composition is U_r: 60%, 72% H_2, 18% CO_2, 10% H_2O) and for the cathode, CO_2 enriched air (14% O_2, 30% CO_2, 56% N_2). The gases are fed and a cell current of 7.5×10^2 A and a cell voltage of 0.850 V are attained. Calculate the cell voltage assuming that it is determined by the utilization of every reactive gas. Use $\Delta G° = -195.8$ kJ/mol, $\Delta H° = -246.1$ kJ/mol. (Numata)

6. Review the superior properties of the proton exchange membrane cell in respect to its use as a power source in vehicles, (a) On what fuel will it operate? (b) What is the storage medium for such a fuel and how is the storer re-formed to give H_2? (c)Assuming as a simplification that planetary warming is entirely due to automotive CO_2 production from fossil fuel-driven vehicles, and that the fuel cell efficiency of conversion to electricity is twice that resulting from internal combustion of fossil fuels, what would be the decrease (at a constant world population of emitting vehicles) in the rate of emissions of CO_2 if all cars were run on H_2 in fuel cells formed from on-board re-forming from methanol? (Bockris)

7. (a) Calculate the amount of hydrogen and oxygen needed to operate a PEM fuel cell for 1 hr, producing a power of 10 W at 0.75 V from a single cell of an area of 25 cm^2. Assume 100% utilization of reactants. (b) What is the voltage efficiency of the cell? (Dhar)

8. One 18-cell stack of a 50-cm^2 area per cell produces a power of about 270 W at 13.5 V when operating with hydrogen and oxygen at atmospheric pressure and 70 °C. The same stack produces a power of about 175 W at 13.5 V when operating with hydrogen and air. Calculate the amount of hydrogen, air, and oxygen needed to operate the stack for 4 hr. Assume 100% utilization of hydrogen and oxygen, and 100, 75, and 50% utilization of air. (Dhar)

9. One3-kW 45-cell stack for providing power to a household has an area of 400 cm^2 per cell. The stack is to be operated with re-formed hydrogen from propane as the fuel and air as the oxidant. Propane breaks down in the presence of air to an approximate composition of 21% H_2, 14% CO_2, and 65% N_2. (a) How much of the re-formed fuel would be required at 90% efficiency to run the stack for 1 hr at 28 V? (b) How much air would be required at 50% efficiency? (c) How much propane would be required? Assume the propane breakdown reaction is as follows: $CH_3-CH_2-CH_3 + 3O_2 \rightarrow 4H_2 + 3CO_2$. (Dhar)

10. In a PEM fuel cell, hydrogen is oxidized at the anode and oxygen is reduced at the cathode. The overall reaction is

$$H_2 + \tfrac{1}{2} O_2 \rightarrow H_2O$$
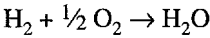
You are operating a 1.2-kW fuel cell stack (20 cells) by using H_2 and air. Assume the utilization for air is 50% and the output voltage of the stack is 12V. The air contains 20% O_2 and is at 25°C and 1 atm. Estimate the consumption rate of air (liters/hour). (Ho)

11. The open-circuit emf of an H_2-consuming solid oxide fuel cell at 900°C was measured to be 1.02 V. (a) If p_{O_2} = 100 kPa at the cathode, estimate the p_{H_2}/p_{H_2O} ratio at the anode.
    By varying the external resistive load, R_{ex}, of a cell, the following current densities were obtained at the corresponding potentials (E):

TABLE E.1

$J/mA/cm^2$	E
0	1.02
150	0.9
305	0.8
455	0.7
750	0.5
920	0.4
1060	0.3
1205	0.2
1340	0.1
1500	0

(b) Find the thermodynamic efficiency of each operating potential. (c) Plot the data and find the maximum power output of the cell. (d) What type of overpotential do the data suggest? (e) Evaluate the area specific resistance of the cell in ohms per square centimeter. (Ho)

12. Calculate the free energy change (heat change) of the cell reaction (ΔH) in calories for two battery systems: (a) A lead-acid cell with an open-circuit voltage of 2.01 V at 15 °C and a temperature coefficient of resistance (dE/dT) of 0.0037 V/K. (b) A Zn-Hg cell (Clark cell) with an open-circuit potential of 1.4324 V at 15 °C and a temperature coefficient of 0.00019 V/K. (Bhardwaj)

13. A criterion for judging the degree of charge of a lead-acid battery

$$-Pb|PbSO_4|H_2SO_4|PbSO_4|PbO_2|Pb +$$

is the density ρ of its electrolyte (concentrated sulfuric acid). A battery is usually considered fully charged when its electrolyte is 40% H_2SO_4 ($\rho = 1.3$ kg dm^{-3} at 25 °C) and discharged when its electrolyte is 16% H_2SO_4 ($\rho = 1.1$ kg dm^{-3} at 25 °C). Determine the cell emf at the beginning and at the end of the discharge process at 25 °C.

TABLE E.2

$E^\circ_{PbO_2|PbSO_4} = 1.682$ V; $E^\circ_{PbSO_4|Pb} = -0.3588$ V; mean activity coefficients and osmotic coefficients for H_2SO_4 are as follows:

$M_{H_2SO_4}$	0.5	0.7	1	1.5	2	2.5
$\gamma_\pm$	0.1557	0.1417	0.1316	0.1263	0.1276	0.1331
ϕ	0.676	0.689	0.721	0.78	0.846	0.916
$M_{H_2SO_4}$	3	3.5	4	4.5	5	5.5
$\gamma_\pm$	0.1422	0.1547	0.17	0.1875	0.2081	0.2132
ϕ	0.991	1.071	1.15	1.226	1.303	1.376
$M_{H_2SO_4}$	6	6.5	7	7.5	8	
$\gamma_\pm$	0.2567	0.2852	0.3166	0.35	0.386	
ϕ	1.445	1.512	1.576	1.636	1.691	

14. Zinc is the most popular negative pole for commercial primary alkaline cells. When building such cells, one must necessarily use high-purity zinc and/or additives like mercury (now unpopular for its ecological impact). (a) Can you guess why?

 Zinc can be combined with many positive pole partners. Consider the following three examples:

$$Zn \mid ZnO + KOH \mid MnOOH \mid MnO_2$$

$$Zn \mid ZnO + KOH \mid Ag \mid Ag_2O$$

$$Zn \mid ZnO + KOH \mid Hg \mid HgO$$
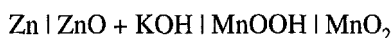
 (b) Write the relevant cell reactions and discuss the advantages of each configuration. (Mussini)

15. A battery-operated radio uses an AA-size alkaline manganese dioxide battery. If the radio drains 16 mA current at average voltage of 1.2 V for 10,140 min from the battery and the battery dies completely, calculate (a) the total energy stored in the AA battery and (b) the load resistance of the radio in ohms. (c) If the same battery is used in an audio cassette player that uses 120 mA/hr, how long will this battery last? (Bhardwaj)

16. In a 60 A hr lead-acid battery, sulfuric acid would be consumed during the discharge. The specific mass of the H_2SO_4 solution is 1.25 at a fully charged

state and decreases to 1.1 at a fully discharged state. In a fully charged state, the volume of H_2SO_4 is 810 ml and the concentration is 33.3 wt.%. Calculate the decreased volume (ml) of the H_2SO_4 solution after 60 A hr of full discharge. (Ho)

17. A diagram in the chapter shows the difference between the energy available in the discharge of a battery and that for charging it. Thus, there is an inevitable energy loss in the cycle of discharge and charging in a battery. Quantify the advantage (from the viewpoint of minimizing energy losses) of discharging and charging a hypothetical battery with an open-circuit potential (zero current drain) of 2.00 V at a lower rather than a higher current density. The increase of potential per decade of current density is 0.2 V, and the loss on discharge, the same. The battery is on two charge-discharge cycles, the higher of which is 100 mA cm^{-2} and the lower 10 times less. The plots of the battery voltage on charge and discharge against log current density intersect at 10^{-5} A cm^{-2}. (Bockris)

18. Electric cars working on batteries have zero emissions. Why, then, are major automotive companies planning fuel cell-driven cars emitting CO_2? (Bockris)

PROBLEMS

1. The first practical use of fuel cells was to provide auxiliary power in space vehicles. What was the principal reason for this choice? Considering fuel cells running on H_2 and air O_2, supplied by the re-forming of CH_4 or CH_3OH, discuss the various fuel cell types and suggest what applications will be made of re-formed H. (Bockris)

2. Calculate the area of solar collection necessary to power a 300-kW lighter-than-air vehicle by means of photovoltaic cells (20% efficient), electrolyzing on-board water. Assume a 200-m-long dirigible (as an approximation, take its shape as cylindrical with a radius of 20 m). Could a hydrogen-lifted vehicle be solar powered but fly during darkness? (Bockris)

3. It is often claimed that electrocatalysts in fuel cells are dependent on the exchange current density, i_0 of the slowest reaction in the cell. (a) Make Tafel plots for $i_0 = 10^{-1}$, 10^{-6}, and 10^{-3} A cm^{-2} and $b_{Tafel} = 0.12$. (b) Then draw plots of the same type and the same i_0, but with b values of 0.12, 0.05, 0.038, and 0.029 (T = 298 K). (c) Write out your conclusions concerning the interplay of i_0 and b in the Tafel relation $(B = RT/\alpha F)$. (d) How does this relate to the choice of electrocatalytic surfaces for optimal fuel cell performance? (Bockris)

4. The efficacy of porous electrodes depends on two factors: The pores allow an external geometric electrode area of 1 cm^2 to be increased many times and the meniscus-like structure of the three-phase boundary gently increases the local limiting current density. Assume a model electrode with pores having a uniform

radium of 1 μm and a length of 1 mm. Fifty percent of the electrode is filled with such pores, each of which is uniformly active. (a) Calculate the effective area of a porous electrode with these (extremely idealized) assumptions.

In the above idealized structure, assume a specific conductance of the electrolyte of 3 S cm^{-1}. (b) Calculate the ohmic resistance of a pore 1 mm in length and a radius of 1 μm. (c) What effect would this have (in a better model) on the distribution of current in the pore assuming the reaction is the cathodic reduction of O_2 and the catalyst is uniformly distributed?

A cathodically polarized air electrode (planar) has a limiting current of about 10^{-4} A cm^{-2}. In a fuel cell and the critical quantity that controls the magnitude of the current density is the thickness of the electrolyte in the meniscus of the three-phase boundaries. This varies with the shape of the meniscus and the contact angle. (d) Assume a section of the meniscus having a solution thickness of 10^{-5} cm and calculate the limiting current of this section. (e) In light of these zeroth approximation calculations, where do you think the maximum activity of a pore lies? (Bockris)

5. The key point about the design of the monolithic fuel cell is that the fuel gas and its partner (usually air) pass through passages in the cell and meet the electrolyte with constant concentration throughout the gas passages. Thus, each pore is fully active (contrast the situation in most fuel cell pores). In principle, the length of the fuel cell passage is unlimited. (a) Make very rough calculations as to the size of such a cell needed to power a large office building in which there are 5000 employees consuming $1\frac{1}{2}$ kW per person levelized power for 8 hr/day. (b) Make corresponding rough estimates of the cell weight needed to power a vehicle using 100 kW.

Notes: The cells are made of ceramic materials, the density of which can be taken at 2.5. The potential difference across the terminals is 0.7 V. The layers are 50 μm thick. The effective surface area per unit volume is 10 cm^2 cm^{-3}. IR drop = 0.05 V at a current density of 1 A per apparent square centimeter. (Bockris)

6. The thermodynamic efficiency, ε, of a fuel cell is defined from:

$$\varepsilon = \frac{W}{-\Delta H^\circ}$$

where W is the electrical work produced and ΔH° is the enthalpy change in the overall chemical reaction taking place in the fuel cell. (a) Show that ε can be expressed as

$$\varepsilon = \frac{E}{E_{th}} = \frac{-\Delta G^\circ}{-\Delta H^\circ}\frac{E}{E_{rev}}$$

where E is the operating fuel cell potential, the thermoneutral potential E_{th} is defined from $E_{th} = -\Delta H°/nF$, and the open-circuit or reversible potential E_{rev} equals $-\Delta G°/nF$. (b) Show that when $E = E_{th}$, the cell does not produce or consume heat. (c) Under what conditions can ε exceed unity? (d) Find specific examples among the following overall reactions taking place in a solid oxide fuel cell at 800 °C:

$$H_2 + 1/2\ O_2 \rightarrow H_2O$$

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$$

$$NH_3 + (5/4)O_2 \rightarrow NO + (3/2)H_2O$$

$$CH_4 + 1/2O_2 \rightarrow CO + 2H_2$$
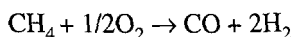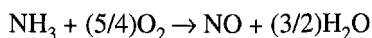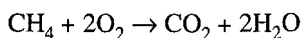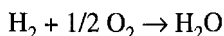
(e) What is a necessary requirement regarding $\Delta S°$ for ε to exceed unity? (f) Analyze a fuel cell operating with $\varepsilon > 1$ from the point of view of the first and second laws of thermodynamics. (g) Does the cell produce or absorb heat? (Vayenas)

7. In this problem the objective is the functioning of a hydrogen-oxygen fuel cell in phosphoric acid at 180 °C and a gas pressure of 8 bar. The following are in kJ/mol at 180 °C, for the combustion of hydrogen with oxygen:

$$\Delta G_{180} = -219; \quad \Delta H_{180} = -244$$

The individual cell consists of two electrodes of dimensions 1.20×0.80 m^2 separated by a ceramic membrane of 1 mm thickness that contains acid having a conduction of $\sigma = 0.1\ \Omega^{-1}\ cm^{-1}$.

Thermodynamic Study. (a) Calculate the electrode potentials and the emf of the cell at equilibrium. (b) Calculate the theoretical value of the reversible potential at 180 °C. (c) Compare these results with that of a heat engine functioning between 180 and 480 °C.

Kinetic Study. (d) Evaluate the resistance of the electrolyte R_e. (e) Calculate the diffusional resistance R_d associated with a limiting current of 5 A/cm^2 for the hydrogen electrode and at 1 A/cm^2 for the oxygen electrode. (f) Evaluate the overpotentials for the charge transfer, η_1, for a current density, $i = 400$ mA cm^{-2}. The exchange current densities given are 200 mA m^{-2} for hydrogen, $i_{0c} = 1$ mA m^{-2} for the oxygen. One can assume that the transfer coefficients are 0.5 for both reactions. (g) Calculate the emf in the operating conditions and the power density in watts per square centimeter. (Lamy)

8. The potential difference of an ideal reversible electrochemical cell in open circuit is 0.965 V at 25 °C and 1 atm. The open-circuit potential was measured

at different temperatures and the relationship between these two variables was found to be:

$$E = 0.924 + 0.0015t + 6.1 \times 10^{-6}t^2$$

where E is the open-circuit potential in volts and t is the temperature of the cell in degrees Celsius. Electrical energy is supplied to the cell in a reversible form using an external source until the charge is equal to the faraday constant. (a) How much heat was exchanged between the cell and a thermostat during this process? Consider for simplicity that electron exchange is equal to one, and that the thermostat works at 298 K. (b) If the cell is short-circuited in such a way that no electrical work is obtained and the cell is allowed to return to its initial conditions, how much heat is exchanged between the cell and the thermostat? (Zinola)

9. The emf of a cell or battery depends on the concentration of electrolyte. Calculate the emf of a charged and discharged lead-acid cell that has 29% by weight of sulfuric acid when it is fully charged. On discharge of this cell, the concentration of acid reduces to 21 % by weight. Assume that the temperature is 25 °C and the standard potential ($E°$) for both concentration is 2.0359 V. (Bhardwaj)

10. A car company wants to design a battery for their hybrid vehicle. The battery requirement for this system is 300 V with a nominal capacity of 3.0 kW hr. Suppose they have a 5 A hr cell at 2.0 V. (a) How many cells will be required to make the required battery? (b) How many cells will be in series and how many in parallel? (Bhardwaj)

11. Calculate the theoretical energy density of a lead-acid battery at 25 °C. Assume that 1 mol each of lead and lead dioxide is discharged from an initial H_2SO_4 concentration to final acid concentration. The 54 A hr are produced at room temperature in this discharge at an average voltage of 2 V. Base your calculations only on moles of the three active materials, i.e., lead, lead dioxide, and sulfuric acid. (Bhardwaj)

12. A 2.16-V battery is discharged at a constant current of 5, 40, and 120 A. The average voltages during discharge at these currents are 2.0, 1.9, and 1.87 V, respectively. The run time for the battery is 3600, 370, and 90 sec, respectively, at these currents. The weight of this battery is 460 g. (a) Calculate the power, specific power, and specific energy of this battery in watts, watts per kilogram, and watt hours per kilogram at 5, 40, and 120 A constant current. (Bhardwaj)

13. A silver-zinc battery has a high and steady current density output during operation and has been used widely. The overall reaction is

$$Ag_2O_2 + 2Zn + 2H_2O \rightarrow 2Ag + 2Zn(OH)_2$$
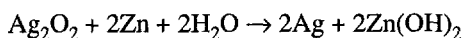
Now we want to design a silver-zinc battery with a capacity of 60 A hr. A 40% KOH solution is used as the electrolyte in the battery, i.e., water consumed during discharge is provided from the electrolyte. Calculate the minimum amount (g) of 40% KOH solution required in this battery. (Ho)

14. (a) Why was Li regarded for so many years as a substance that could not be used in batteries, although the amount of electricity stored for unit weight is greater than in any other substance (except H which, in the form of H_2, needs a weighty container)? (b) What caused the paradigm shift? (c) What does "intercalation" mean? (d) Use examples in your explanation. (e) Discuss the reduction reaction in the counter-electrode to Li. (f) Does its complexity overshadow the possible use of a rechargeable Li battery? (g) What about leaks of H_2O and O_2 from the air in a (so-called) sealed Li battery? (h) What would happen to any traces of water that enter? (i) Could this be the Achilles heal of the Li battery? (Bockris)

15. The Zn-MnO_2 battery is considered "a primary," i.e., not rechargeable. Research shows that the difficulty of recharging it occurs in processes involving use of the *second* electron. (See the text.) (a) Considering a 20-g quantity of Zn, all of which is to be used in the battery action, calculate the weight of MnO_2 that will be just enough for a 1e use of the Zn (i.e., the MnO_2 does not undergo full reduction upon discharge). The battery thereby becomes 1e rechargeable, at half the electrical capacity for the full 2e cell. (b) Use the equation in the text concerning the reactions of this battery. (Note that in reality, not all the MnO_2 and Zn are physically available. Further, the MnO_2 is mixed with conducting carbon, etc.) (Bockris)

16. Capacitors have two great advantages over electrochemical batteries as energy storers: Their recharge time is negligible and they can be recharged an indefinitely large number of times. However, such storers also carry a heavy burden. The energy density that they carry is only 0.1 1 W hr kg^{-1}. Electrochemical capacitors are much better, but still give only 1 W hr kg^{-1}. (a) Discuss in a way that amplifies the text, a 20-year future in the possibilities for these electrochemical devices. (b) What of the possibility of porous structures? Consider an RuO_2 structure with a real to apparent surface area of 500. (c) What would be the watts per kilogram here? (Bockris)

17. Ragone plots reveal some characteristic differences of fuel cells and batteries. To put the matter succinctly, batteries are known for their power and fuel cells for their energy, both per unit of weight. Discuss these characteristics as the basis for hybrid designs in the powering of automobiles. (Bockris)

18. The aluminum-air battery produces electricity by means of a 3e dissolution of Al in an alkaline medium. The $Al(OH)_3$ produced is to be recycled to Al. Consider the volume of the gas tank of a car to be that equivalent to 18 gallons of gasoline and that half of this space is occupied by Al in foil form which

unwinds and is fed into the battery. Assume a cell potential of 1.0 V at 100 mA cm^{-2}. What will be the range of a 25-kW car powered in this way?

Note that the manufacture of Al occurs from plentifully available natural resources; at present this is bauxite and eventually it will be clay. The electricity for many aluminum plants comes from hydroelectric plants, so that the use of mechanically rechargeable Al batteries (obtained in this manner) would lead to a reduction in greenhouse gases faster than that of re-forming methanol on-board cars to obtain hydrogen for fuel cells (because this gives rise to CO_2). (Bockris)

19. Professor S. Srinivasan and his team have studied the effect of pressure and characteristics of the current–potential relations in a hydrogen-oxygen fuel cell with a proton exchange membrane (Y. W. Rho, O. A. Velev, S. Srinivasan, and Y. T. Kho, *J. Electrochem. Soc.* **141**: 2084, 2089, 1994). In this problem, it is proposed to study the applicability of the theoretical dependence of the cell potential as a function of pressure. The temperature is 25 °C and it may be assumed that the pressure of the gas in each of the compartments, i.e., the anodic compartment (hydrogen) and the cathodic compartment (oxygen), are the same, $p_{H_2} = p_{O_2} = p$. For the formation of water in its standard state, the relevant thermodynamic quantities are:

$$\Delta G^\circ = -237 \text{ kJ/mol}$$

and

$$\Delta H^\circ = -285.8 \text{ kJ/mol}$$

(a) Write the electrode reactions and the overall reactions, (b) Calculate the standard emf, E°_{rev}. (c) Write the electrode potentials in the reversible condition at zero current for the anode, E°_{an}, and for the cathode, E°_{cath}, as a function of pressure of the gases p_{H_2} and p_{O_2} and of the standard pressure of $p_0 = 1$ atm. (d) Deduce the reversible potential at the cell, (e) Express the electrode potential E at a current density of i. Assume that the reaction works with the hydrogen oxidation in the reversible region so that the potential–current relation is linear. For the reduction of oxygen, one can apply the Tafel law with $i_0 = k[p_{O_2}]^{5/4}$ using a value of $\alpha_{cath} = 1$, where α_{cath} is the cathodic transfer factor. (f) Deduce an equation for the characteristic current density–potential relation of the elementary cell in the form:

$$E = E_0 - b \log i - R_i,$$

where $R = R_t + R_e$ and $E_0 = E^o_{rev} + b \log i_0 + c \ln p/p_0$. In these equations, E^o_{rev} represents the reversible potential of the cell in its standard condition ($T = 25$ °C, $p_0 = 1$ atm), b is the Tafel slope for the reduction of oxygen, c is the equivalent of a Tafel slope for the effect of pressure, and R is the sum of the electrolyte resistance and the resistance to the electron transfer at the interface. (g) Deduce a theoretical expression for c. (h) Using the table of results for E(mV) as a function of i (mA/cm^2) for various values of the pressure p (in atmospheres), draw a curve between E as a function of p where p is the pressure.

TABLE P.1

i	1	10	20	30	40	60	100
$p = 1$	959	900	882	871	862	849	831
$p = 3$	1014	955	937	926	917	904	886
$p = 5$	1040	981	963	952	943	930	912
$p = 10$	1075	1016	998	987	978	965	947
i	200	300	400	600	800	1000	
$p = 1$	801	779	759	724	691	660	
$p = 3$	856	834	814	779	746	715	
$p = 5$	882	860	840	805	772	741	
$p = 10$	917	897	875	840	807	776	

(i) Calculate the resistance of the electrolyte and the exchange current density i_0. (j) Deduce the value of the coefficient c, where $c = \delta E/\delta \ln p$, which is well verified by the results ($R = 8314$ J mol^{-1}, $F = 96{,}490$ C, and 0°C = 273.15 K). (Lamy)

20. In a solid electrolyte fuel cell utilizing YSZ (an O^{2-} conductor), as the solid electrolyte, H_2 and O_2 are consumed at the anode and cathode, respectively, according to the reactions:

$$H_2 + O^{2-} \rightarrow H_2O + 2e^-$$

$$\tfrac{1}{2} O_2 + 2e^- \rightarrow O^{2-}$$
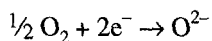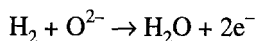
Evaluate the open-circuit emf of the cell at 800, 900, and 1000 °C when at the anode $p_{H_2} = 80$ kPa, $p_{H_2O} = 20$ kPa, and at the cathode $p_{O_2} = 100$ kPa. (Makri, Pitsellis, and Vayenas)

21. The open-circuit emf of an H_2-consuming solid oxide fuel cell at 900 °C was measured to be 1.02 V. (a) If $p_{O_2} = 100$ kPa at the cathode, estimate the p_{H_2}/p_{H_2O} ratio at the anode.
    By varying the external resistive load, R_{ex}, of the cell, the following current densities were obtained at the corresponding potential E:

TABLE P.2

$J/mA/cm^2$	E
0	1.02
150	0.9
305	0.8
455	0.7
750	0.5
920	0.4
1060	0.3
1205	0.2
1340	0.1
1500	0

(b) Find the thermodynamic efficiency at each operating potential. (c) Plot the data and find the maximum power output of the cell. (d) What type of overpotential do the data suggest? (e) Evaluate the area-specific resistance of the cell in ohms per square centimeter. (Makri, Pitsellis and Vayenas)

MICRO RESEARCH PROBLEM

1. Consider a manufacturing community of 100,000 persons, the levelized total energy consumption of which (including industry and military) is 10kW/person. Further, assume that one-quarter of these persons are employed in the synthesis of chemicals, using the second fuel cell principle (Section ??). As a simplification, assume the manufacture is carried out by fuel cells working at 0.6 V and a current density of 1 A cm^{-2}. The manufactured item has a molecular weight of 300 and requires 8 e_0 per mole in its synthesis.

   (a) If the manufacturing facility is to supply half of the total energy of the community, how many tons of material will have to be manufactured per month? (b) How much would the manufacturers earn if they sold the by-product electricity at 5¢ kW hr^{-1}? (Bockris)

This page intentionally left blank

CHAPTER 14

BIOELECTROCHEMISTRY

14.1. BIOELECTRODICS

14.1.1. Introduction

Electrochemistry as we have come to know it in the preceding thirteen chapters consists in the study of ionic solutions, and electrodes where ions and electrons combine and separate. It seems a far cry from biology which is, surely, all about flesh and blood and bone. Yet, right from the beginning with Galvani in that laboratory in Bologna, Italy, in 1791, bioelectrochemistry has been a part of electrochemistry. Historically, in fact, electrochemistry came out of it.

Thus, it is not too much to say, as Becker did in the book *The Body Electric,* that there are electrochemical events going on in living systems wherever you pry into them. The nervous system is certainly based on the flow of electric currents and it is not at all fanciful to see nerves as the wires that run between the enzymes, the electrodes of the body. Bodies are full of membranes, too, and so are electrochemical cells. Some reactions in the body baffle chemists by going *up* free energy gradients, but again, that is just what happens in electrolysis, in electrochemical reactors. Then, we can look to the brain, and what is the least-known and documented phenomenon there? Is it not the electrical oscillations that (Maxwell-Cade, 1996) have been shown not only to signal mental activity but to be connected with consciousness itself? They, too, have, an electrochemical origin.

This recital of electrochemistry in the body could go on for some time. However, there are also factors that reminds us of the little we actually know. To imply that we can study electrochemical phenomena in the *immense* complexity of living systems when all we know is how to explain simple systems like fuel cells and corrosion seems to be the crassest arrogance. Only 50 years ago, bioelectrochemistry was still at the Nernst stage, with membrane potentials and formulas such as $(RT/F) \ln (a_{1+\text{out}}/a_{1+\text{in}})$ for the potentials observed. The science of biology is a truly gigantic edifice, so big,

in fact, that it includes all of organic chemistry and uses it to explore very, very complicated interactions. Debye once said that the greatness of a scientist was measured by his ability to know when and how to approximate. One certainly needs to exercise this aspect of one's work when it comes to bioelectrochemistry for here the jump from reality to what we electrochemists can conceptualize is particularly great—far greater than with electrochemical energy storage or the electrosynthesis of nylon.

Nevertheless, we are bound to try. The reason is not only because the electrical aspects of physiology are evident everywhere, but because bioelectrochemistry may lead us across the great divide between what we know from our materialistic, reductive science of the twentieth century and what we don't understand at all. There are phenomena, not a few in number and very real, that scientists seldom like to recognize because they have no explanation for them—for example consciousness, telepathy, and strange, inexplicable healings. If bioelectrochemistry leads us to uncover all the undiscovered country that is accessible to the open mind, then our small efforts to explore this immense field are justified.

14.1.2. Useful Preliminaries

14.1.2.1. *Size.* Up until now in this book, the ions and molecules that have been actors on our stage have been exemplified by H^+, O_2, ZnO, and the metal hydrides or chelating oxides, such as CuO. However, this is an inadequate introduction to the molecules that make up the cells of biological systems. The amino acid, glycine (i.e., $NH_3^+-CH_2-COO^-$) would be the simplest example (excluding of course H_3O^+, H_2O, and O_2) of an entity that takes part in bioreactions. A structural element within the amino acids is the peptide group

$$-N-C-$$
$$| \quad |$$
$$H \quad O$$

which is important because when many of these groups occur in a chain, and such chains form a polymer, one has a protein. Proteins form skin, nails, and skeletal structures. Enzymes, biocatalysts, are proteins. Hemoglobin, which carries O_2 around the body, is a protein. Proteins can have molecular weights as low as 10,000 but some are really very large, with molecular weights of 50×10^6. The corresponding radii of the larger of these entities (if they were formed spherically), would be in the hundreds of angstroms. Examples of some amino acids are shown in Fig. 14.1. Structures that would form a protein would be, for example, glycine.

The proteins found in nature are made up of only about 20 different, individual amino acids. On the other hand, a typical protein consists of several hundred of these 20 distinguishable amino acids. If one calculates the number of ways in which, say, 500 things, 20 of which are different entities, can be arranged, the answer is
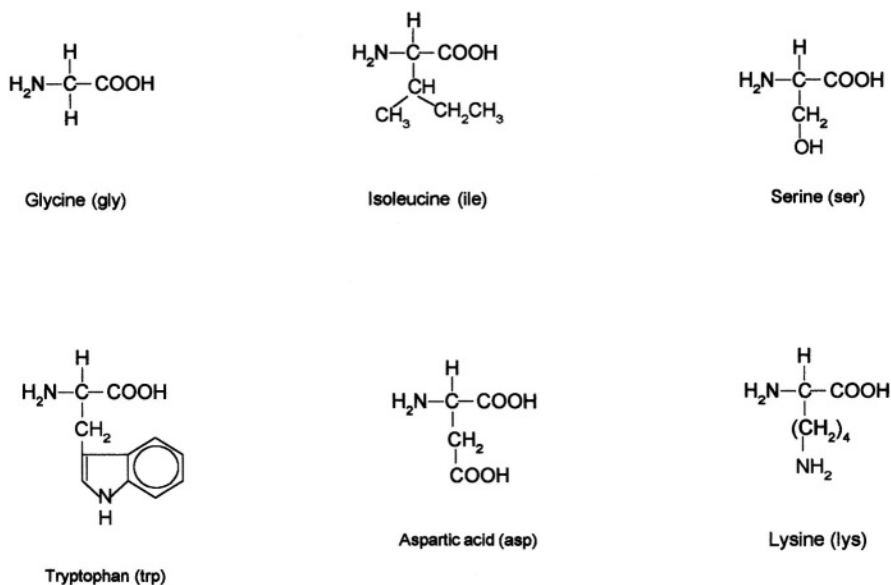
Fig. 14.1. Examples of some amino acids. (Reprinted with permission from Accounts of Chemical Research **21** copyright 1989, Fig. 1, American Chemical Society.)

$$\frac{500!}{20!}$$

which is about 10^{1227}. Some cosmologists have presented calculations for the number of particles in the universe as being in the region of 10^{93}.

These proteins (containing several hundred of the 20 special amino acids) might be thought at first to be long, long chains containing repeated peptide groups (see above), which can be written more explicitly as
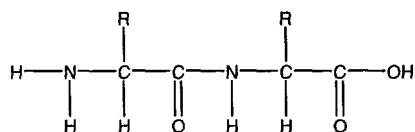
DIPEPTIDE

where the R's may be hydrocarbon chains on other peptide elements and the shaded area denotes the bond between the two peptides. However, these long, flexible chains are neither linear nor random in shape. They coil and stretch in a way that greatly affects the properties of the protein—how it does its work. Among the most important of these structures is an arrangement called the α-helix (Pauling and Corey, 1951), but
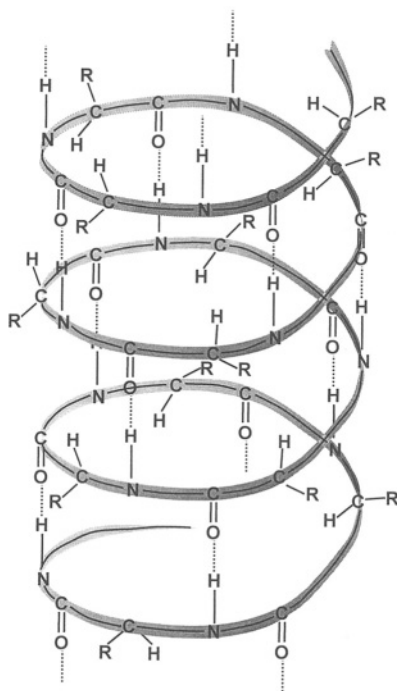
Fig. 14.2. The α-helix structure for a protein. (Reprinted from T. L. Brown and H. E. LeMay, *Chemistry: The Central Science*, Prentice-Hall, 1977.)

it would take us too far away from our main themes to describe its chemical properties (Fig. 14.2).

The great importance of this structure is that it was the forerunner to the elucidation of the structure of DNA, deoxyribonucleic acid, in which two α-helix coils are joined together in a double-stranded helical structure containing a number of units of a sugar called deoxyribose, to which is attached one of four organic bases (Fig. 14.3). We must forgo further material on the structure and implications of this very remarkable and important protein, the molecular weight of which is in the region of 10^6. Its discovery by Crick and Watson at the University of Cambridge in the United Kingdom in 1953 is regarded as one of the greater discoveries in the history of science because the structure occurs in various forms in the cells of all living organisms. Its detailed structure is species specific.[1]

---

[1]Elementary accounts of these discoveries can be found in modern textbooks of general chemistry and, of course, in modern texts on molecular biology.
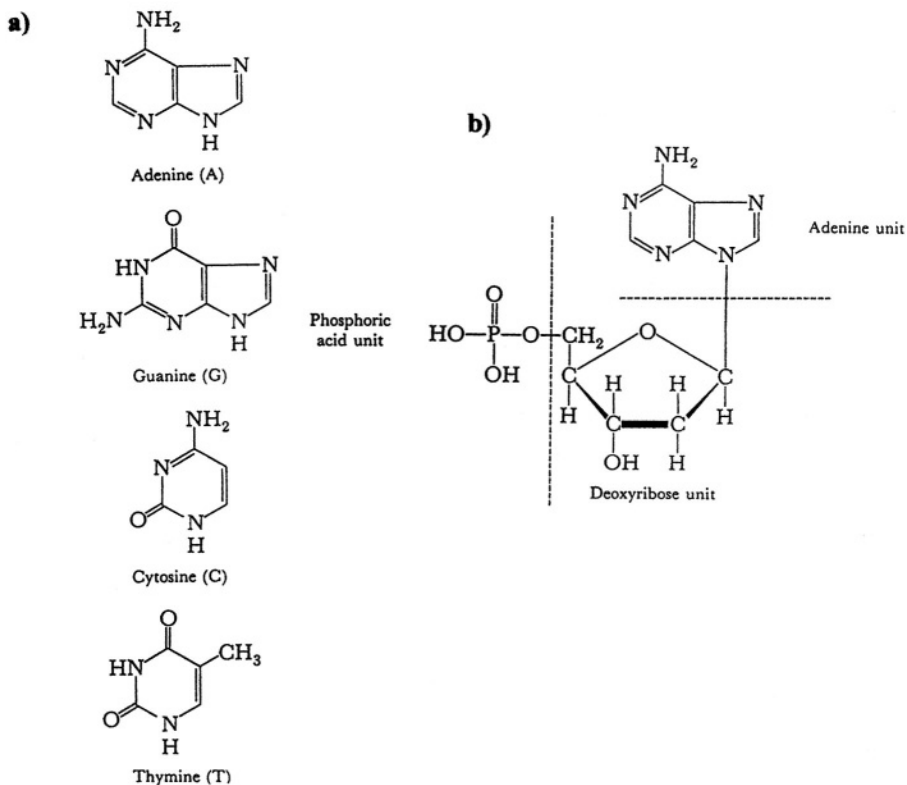
Fig. 14.3. (a) The four organic bases present in DNA. Each base is attached to a ribose molecule through a bond from the nitrogen. (b) Structure of deoxyadenosine, a nucleotide formed from phosphoric acid, deoxyribose, and an organic base, adenine. (Reprinted from T. L. Brown and H. E. LeMay, *Chemistry: The Central Science*, Prentice-Hall, 1977.)

14.1.3. Why Should Electrochemists Be Interested in Amino Acids, Proteins, and DNA?

The first reason is the cultural one. Humankind is passing into a new era, one in which it can consciously control the species living on the planet. Thus, every scientist should know at least the building blocks that will be used in the genetic engineering of the twenty-first century.

However, there is a more immediate reason. There is much electrochemistry mixed in with all the rest of the chemistry controlling the behavior of biomolecules. Thus, the peptide groups that make up proteins can be charged (e.g., NH_2 can form NH_3^+ in acid solution, and COOH tends to ionize to form $COO^- + H_3O^+$) when the pH

is high. The effects arising from the electrical charges in these very large structures are similar to those of the ionic atmosphere discussed in Chapter 3. They give rise to Coulombic attraction and repulsion and dispersive attractions; the larger the entities become, the more they behave as colloidal particles, with all the electrical interaction this implies (Section 6.10.2).

Moreover, intramolecular interactions within the protein groups may increase at extremes of pH so much that the α-helix and the double helix of DNA experience so much internal repulsion that the coil configuration becomes unstable and undergoes what is termed "denaturation," a destructive change resulting in a tangled, disordered structure ("a statistical coil"). A similar change occurs when living organisms are heated to a certain degree above their body temperatures.

Another electrochemical aspect of biostructures, here briefly reviewed, is the need for their stabilization in solution by means of hydration. A schematic of hydration in a protein was shown in Chapter 2 (Fig 2.4). On average, about 20 water molecules are needed to bring stability to each structural unit of nucleic acid.

Thus, the stability of DNA and other biostructures is intimately bound up with electric fields and the electrochemical properties of ions. In these properties lies the explanation of many biophysical phenomena.

14.1.4. Cells, Membranes, and Mitochondria

Robert Hooke (of Hooke's law) was the first to discern (in 1665) a central fact of biology, namely, that biological structures consist of honeycomblike structures, which he called cells. In fact, such cells are to biological systems what crystallites are to metals or sand to beaches.

By the middle of the twentieth century, the constituents of cells were known and could be seen in electron microscopes. Each cell contains a number of structures (organelles), the names and approximate shapes of which are shown in Fig. 14.4.

The question of what surrounds the liquid inside cells, with all the objects floating in it, was clarified by the 1920s. Cells are contained in what might rather loosely be called bags. These bags, membranes, turned out (Gordon and Grendell, 1925) to consist of a layer of phospholipids.[2] Some membranes have single walls and some double. As time went on, this structure was found to be a universal thing; living organisms consist of cells and the cells themselves are held inside membranes.

From the century's beginning, through its midpoint, the electrochemistry of electrodes was based upon the treatment given by Nernst (Section 7.2.36). This had been derived first, for an interface between a metal and its ions in solution, but the treatment had spread (Planck and Henderson, 1890–1907) to the potential difference between two liquids containing different concentrations of electrolytes. The first of these two treatments yields an equation (Nernst equation) identical in form to the

---

[2]These are derivatives of phosphoric acid, in which there is a polar head containing, e.g., NH_3^+ and a long structure based on glycerol but containing 15–16 carbon atom long hydrocarbon groups.
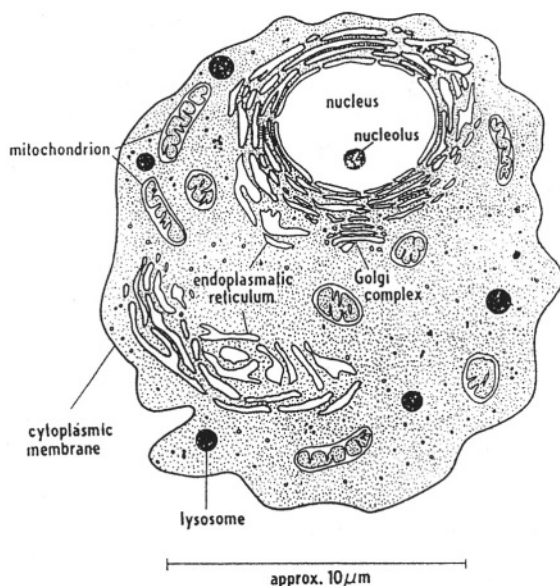
Fig. 14.4. The scheme of an animal cell. (Reprinted from J. Koryta, Ions, *Electrodes and Membranes* Fig. 69. Copyright © J. Wiley & Sons, Ltd. 1991. Reproduced with permission of J. Wiley & Sons, Ltd.)

well-known equation derived in thermodynamics for the electrical potential difference between the electrodes of Glavanic cells:

$$^M\!\varDelta^S\phi = \frac{RT}{F}\ln\frac{c_2}{c_1}$$

where the c's refer to concentration of an entity in the solution and on the electrode, and in the second (Sec. 4.5.9),

$$^{S_1}\!\varDelta^{S_2}\phi = (t_+ - t_-)\frac{RT}{F}\ln\frac{c_2}{c_1} \tag{4.290}$$

t_+ and t_- represent the transport number of cations and anions, respectively. Taking the difference in concentration of the potential-determining entities (2 and 1) as 10 times and $t_+ - t_-$ as 0.15, the two equations yield about 57 and 9 mV respectively and cover very approximately the normal range of the measured values of potential across membranes.

The electrochemistry of membranes will be discussed in the next section (14.2). For now, there is one more thing to note and that is the nature and function of just one

of the elements in cells, the mitochondrion. It has been known for about a generation that mitochondria are the entities in cells where energy is made from the oxidation of organics derived from intake of food and oxygen. Much more will be said about how they may function later (Section 14.8), but the fact is that, as suggested by J. Koryta in 1991, the mitochondria are well described as floating power stations.

14.2. MEMBRANE POTENTIALS

14.2.1. Preliminary

The measurement and interpretation of the potentials across biological membranes has been going on for about a century. A remarkably prescient suggestion was that of Bois-Reymond (1868), who put forward the concept that a cell surface could well be looked at as though it were an electrode.

The principal elements in biological cells are shown in Fig. 14.4. A section of a yeast cell with its membrane is shown in Fig. 14.5. It can be appreciated from these
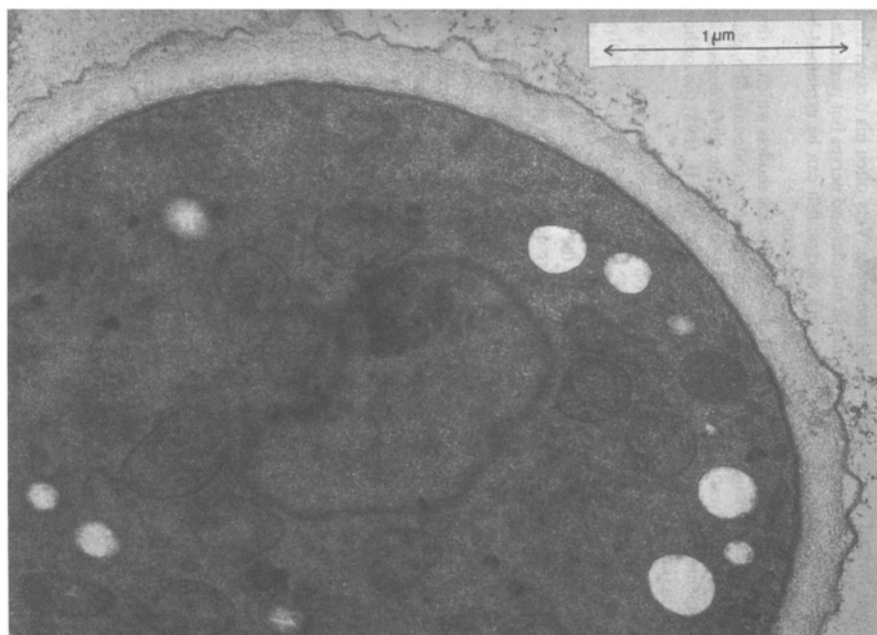
Fig. 14.5. Electron micrograph of a section from a yeast cell. The outer envelope is the cell wall. The inner double line is the cytoplasmatic membrane. (Reprinted from J. Koryta, *Ions, Electrodes and Membranes*, Fig. 70. Copyright © J. Wiley & Sons, Ltd. 1991. Reproduced with permission of J. Wiley & Sons, Ltd.)

two figures that animal cells are rather complex, each one containing the hereditary material, and in particular the entities known as mitochondria, the energy-producing properties of which will be discussed in Section 14.8.

Membranes, which are the subject of this section, can be relatively thick (0.1 mm) if made chemically (see their use in the PEM fuel cell, (Section 13.7.3). Biological membranes are very much thinner (50–100 Å), of the same (3–5 nm) range as that of passive oxides (Section 12.5). Of what do biological membranes consist? Figure 14.6 shows the essential constituents. They are lipids and proteins. How much there is of one and how much of the other varies widely. Thus, in a myelin membrane the lipid content is 80% while at the other end of the range, in mitochondria, there is an inner membrane containing only about 20% lipid. There are many kinds of lipids (as well as very many kinds of proteins), but those in membranes are usually phospholipids and are represented in Fig. 14.7. The structure often contains an H atom and this allows
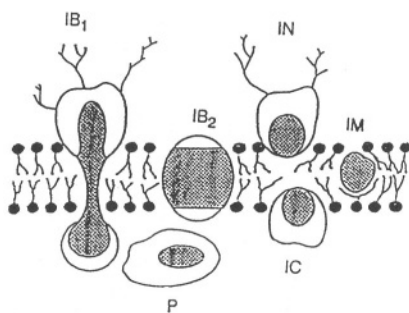
Fig. 14.6. Typical biological membrane structures. A liquid-mosaic model in the form proposed by A. Kortya with different types of disposition of membrane protein) (phospholipids are shown as dark circles with two wavy tails); IB_1, integral, membrane-bridging protein, with a single polypeptide span; IB_2, the same with several spans; IN and IC, integral noncytoplasmic and cytoplasmic proteins; IM, integral buried proteins; P, peripheral protein. (Reprinted from J. Koryta, *Ions, Electrodes and Membranes*, Fig. 81. Copyright © Ltd. 1991. Reproduced with permission of J. Wiley & Sons, Ltd.)
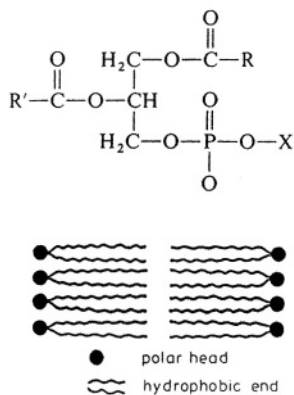
Fig. 14.7. A scheme of the bilayer lipid membrane. The black circles indicate the polar heads (the hydrophilic part) consisting of phosphoric acid, ethanol amine, and analogue derivatives. The wavy lines are the long alkyl chains of fatty acids (the hydrophobic part) (Reprinted from J. Koryta, *Ions, Electrodes and Membranes*, Fig. 83. Copyright © J. Wiley & Sons, Ltd. 1991. Reproduced with permission of J. Wiley & Sons, Ltd.)

the phosphoric acid element to ionize. In the membrane structure, alkyl groups (R and R') are directed inward while the polar groups are on the surface ("fixed charges").

Although many measurements of potentials have been made with membranes obtained from animals, one needs simplification[3] if one is to understand the function of various entities of a cell. The most common model system to act as a simplified biological membrane is the "bilayer lipid membrane" (BLM), first prepared by Mueller in 1962. It consists of two lipid molecules tail to tail (Fig. 14.8) with the polar groups

---

[3]See Debye's statement about the vital importance of being able to approximate—and do it with intuition—as to how far one may go.
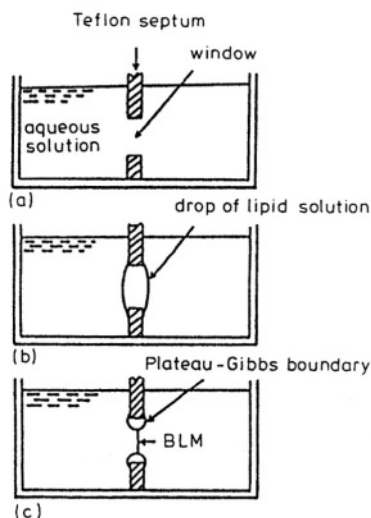
Fig. 14.8. Preparation of a bilayer lipid membrane. (Reprinted from J. Koryta, *Ions, Electrodes and Membranes*, Fig. 84. Copyright © J. Wiley & Sons, Ltd. 1991. Reproduced with permission of J. Wiley & Sons, Ltd.)

oriented to face the solution. In the basic BLM (individual), compounds of a biological system can be built and examined.

These BLMs, although very thin, exhibit very high resistance, up to 10^{10} ohms. Nevertheless, some pores do develop in these membranes and water, followed by ions, enters there and reduces the resistance. Application of a potential increases the flow of ions through the pores and the number of pores; this further reduces the resistance.

The method for measuring a membrane potential is simple. One places an electrode (e.g., a calomel electrode if the solution contains Cl^-) on either side of the membrane, which usually occupies a hole about 1 mm in diameter in a Teflon sheet. Since the potential of the calomel electrode is accurately known and varies according to the Nernst potential with log a_{Cl^-} (Section 7.2.7), the difference in potential arising from the two different Cl^- concentrations on each side of the membrane is easily known and can be subtracted from the total potential differences registered between the two electrodes to give the value due to the membrane. Most of the membrane potentials recorded in the literature lie within values of tens to hundreds of millivolts.

14.2.2. Simplistic Theories of Membrane Potentials

Until the 1950s some bioelectrochemists confidently explained membrane potentials by assuming that only *one* ion (e.g., K^+ in KCl) permeated the membrane. If so, then

$$\bar{\mu}_i^\alpha = \mu + nF\phi^x = \mu_{o,i} + RT \ln a_i^\alpha + nF\phi^\alpha \tag{14.1}$$

would be the expression for the electrochemical potential of the permeable ion, *i,* on the one side, α. On the other side, β, the corresponding expression for the electrochemical potential of the same ion, there would be

$$\bar{\mu}_i^\beta = \mu + nF\phi^\beta = \mu_{o,i} + RT \ln a_i^\beta + nF\phi^\beta \tag{14.2}$$

Now, as long as equilibrium can be assumed between permeating ions, *i,* on each side (e.g., on the α and the β side, respectively).

$$\bar{\mu}_i^\alpha = \bar{\mu}_i^\beta \tag{14.3}$$

Hence,

$$^\alpha \Delta^R \phi = \frac{Rt}{nF} \ln \frac{a_i^\beta}{a_i^\alpha} \tag{14.4}$$

This theory of an equilibrium of *one* species between each side of the membrane was formulated by Donnan in 1925 and from then until 1955, it reigned as *the* theory of membrane potentials. Its demise came when radiotracer measurements showed that all relevant ions (e.g., K^+, Na^+, and Cl^-) permeated more than a dozen actual biological membranes, although each ion had a characteristic permeability coefficient in each membrane (Hodgkin and Keynes, 1953).

Being developed in parallel with the rise and fall of the Donnan equilibrium theory of membrane potentials was the application of liquid junction potential theory to membranes, which the great Nernst himself had proposed as far back as 1888 (see Section 14.1.1). This brought an element of movement into the theory and it became a steady-state theory, rather than one of equilibrium, and involved the mobilities of the various ions. The theory grew by application of the Nernst–Planck equation (4.226) to take into account the driving forces due to concentration and potential gradients, and one form of it was due to Hodgkin and Katz (1949), developed from the 1943 equation of Goldman. This equation marked the end of an era in the study of membrane potentials. The equation is still worth quoting because it was used in the renowned Hodgkin–Huxley theory of the complex electrochemical activities when a nerve fires and sends a message (Section 14.4). It runs:

$$\Delta\phi = \frac{RT}{F} \ln \frac{\Sigma P_i c_{i,\text{in}} + \Sigma P_k c_{k,\text{in}}}{\Sigma P_i c_{i,\text{out}} + \Sigma P_k c_{k,\text{out}}} \tag{14.5}$$

where α and β represent solutions on either side of the membrane, e.g., outer and inner solutions.

The P's are now "permeability coefficients" and are related to the mobilities of the ions as in the original Nernst theory. The subscripts in and out refer to the concentrations of the ions inside and outside the membrane and the P's describe diffusion coefficients, mobilities, and the membrane thickness, but, in the Hodgkin–Huxley theory, were used as adjustable parameters.

14.2.3. Modern Approaches to the Theory of Membrane Potentials

Although theories descended from the liquid junction potential ideas of the nineteenth century have been used to interpret data on the potentials observed across biological membranes, by the 1960s there were a number of substantial objections to the theories as presented. These showed that such theories were not consistent with all the tests of the potentials across biological membranes.

Thus, one can artificially change the concentrations of Na^+, K^+, and Cl^- on either side of the membrane. Then, one can go back to the Hodgkin–Katz equation (14.5) and ask what change in potential these artificial changes of ionic concentrations should bring about. There is found (Jahn, 1962) to be a poor match between theory and experiment. Ionic concentration differences alone, then, do not completely determine membrane potentials in living systems.

Some membrane potentials are affected by light, just as if the membranes were semiconductors. This is entirely outside the capabilities of theories that depend on the interplay of potential and concentration gradients. Are the membranes acting as photoelectrodes (Chapter 10)?

During the 1960s there was a sea change in the views on what was capable of conducting electronically, i.e., be an electrode. In 1949 Albert Szent-Gyorgyi made the seminal suggestion that some biomaterials might be regarded as semiconductors. This was greatly against the paradigm of the time and meant that, instead of being regarded electronically as a pieces of wax, biomaterials could be thought of as possible electronic conductors and hence electrodes. This would be consistent with the idea that the first step in photosynthesis is the photoelectrochemical decomposition of water, and it would account, in principle, for the photosensitivity of membranes.

If Szent-Gyorgyi's concept could be used to interpret the electronic conduction of some biomaterials, there arose the idea that electron transfer at the solid/solution interface could occur with the solid being a biomaterial. Jahn (1962) was the first to come up with the suggestion of a radically different theory of membrane potentials. Instead of seeing this potential as being *caused* by the interplay of concentration and electric gradients in the membrane, he saw these gradients and potentials as *results*—the results of two bioelectrode reactions, a redox reaction occurring on the α side and another on the β side of a membrane. Jahn's concepts pictured the membrane as a bielectrode, with each side the site of differing (but coupled) redox reactions, the membrane itself acting somewhat like the membrane in a fuel cell (Fig. 14.9). A more detailed view, due to Tien, is shown in Fig. 14.10.

But what, it might be asked, of Eq. (14.5), which is derived by simply considering the fluxes of ions across the membrane with never a whisper of electron transfer to and from certain groups on the surface of the membranes? Can a view of the potential difference across biological membranes involving interfacial electron transfer work as well? It might be thought that a radically different view of the membrane potential, engendered by thoughts arising from the introduction of semiconductor properties to biomaterials (i.e., the development of Bois-Reymond's seminal suggestion of 1848 that cell surfaces are like electrodes) would give rise to an equation for a membrane potential utterly different from that originating in the Nernst–Planck concepts behind Eq. (14.5). Could the new ideas related to electron transfer fit the facts?

The difficulty of deducing an equation to represent the "two electrodes back-to-back" of Jahn is less than one might at first think. For we are speaking of two coupled reactions joined by an ionic pathway connecting two hypothetical electrode reactions and occurring at the two surfaces, respectively, of what are regarded as semiconducting materials on which electrons are accepted and rejected (see Figs. 14.9 and 14.10). This is indeed a situation not far different from that of the corrosion couple, and the potential arising in such a situation has already been deduced and is given in Eq. (12.27).

The equations deduced for the two reactions making up a corrosion couple are specialized in the sense that the cathodic reaction is taken to be hydrogen evolution
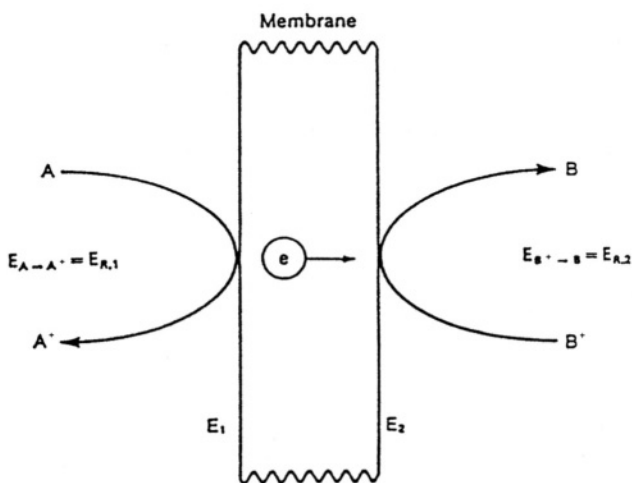
Fig. 14.9. Schematic diagram showing the oxidation and reduction reactions occurring at a membrane solution interface on sides 1 and 2 of the membrane. (Reprinted from M. A. Habibi and O'M. Bockris, *J. Bioelectricity* **2**:66 (1984). Reprinted by permission of Marcel Dekker.)
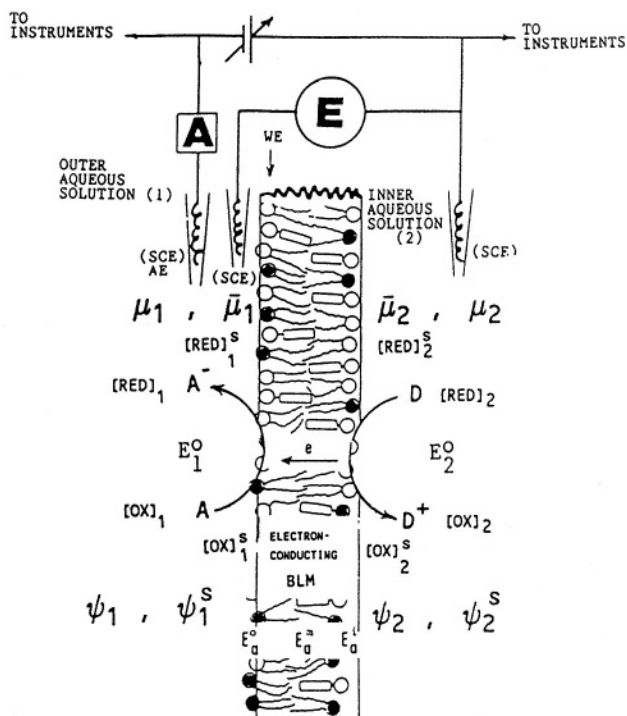
Fig. 14.10. Transmembrane electron movement and redox reactions. Also shown schematically are electrodes and circuit diagram for cyclic voltammetry. WE, working electrode; SCE, saturated calomel electrode; AE, auxiliary electrode. μ, and $\bar{\mu}$ are chemical and electrochemical potentials, respectively. Bulk concentrations of reduced (RED) and oxidized (OX) species on either side of the membrane as indicated by subscripts 1 and 2; interface concentrations are designated by a superscripts (Reprinted from H. T. Tien, *Aspects of Membrane Chemistry*, Kluwer Academic Publishers, 1991.)

and the anodic one, metal dissolution, the α values being assumed to be one-half, but one can drop these assumptions and assume that the membrane potential is simply that of a bielectrode.

Somewhat similar reasoning can be applied to the comings and goings of the Na^+, K^+ and Cl^- ions which (in this view) determine the potential of the membrane. Koryta has formulated this theory by writing electrode kinetic equations similar to those that lie behind the Butler–Volmer equation. Then one can write the sum of the fluxes for positive charges equal to the sum for the charges:

$$J_{Na^+} + J_{K^+} = J_{Cl^-}$$

This procedure results in the following equation for a membrane potential (Koryta, 1991):

$$\Delta\phi_M = \frac{RT}{F} \ln \frac{k_{K^+}^0 c_{K^+}(1) + k_{Na^+}^0 c_{Na^+}(1) + k_{Cl^-}^0 c_{Cl^-}(2)}{k_{K^+}^0 c_{K^+}(2) + k_{Na^+}^0 c_{Na^+}(2) + k_{Cl^-}^0 c_{Cl^-}(1)} \tag{14.6}$$

The model upon which this equation is based neglects potential differences that may occur due to IR *inside* the membrane and assumes that the entire membrane potential consists of the difference of the two interfacial potential differences. In some cases, this may be a good approximation.[4] In other cases, the potential difference through the membrane (determined by fluxes considered in the Nernst–Planck-type equations) may dominate. A comparison of Eq. (14.5) with (14.6) shows that both models lead to equations that have the same form.

What, it may be asked, of radiotracer experiments, demonstrating the movement of ions across the membrane? Of course, such passage of ions occurs. It also occurs across the membranes of working electrochemical cells and in these cases certainly the ions are driven by two interfacial cooperating processes. But such movements occur as results of events elsewhere. The differences in concentration in which they originate are not the *causes* of the potentials with which they are associated. The causes are surface electron-transfer reactions that occur at the interfaces, i.e., at interfaces on either side of the membranes. Such sites of electron transfer are probably at enzymes adsorbed on the membrane surfaces. Thus, in this view, the eventual origin of membrane potentials lies in the free energy changes in redox reactions occurring, respectively, on each of the two membrane surfaces.

14.3. ELECTRICAL CONDUCTION IN BIOLOGICAL ORGANISMS

14.3.1. Electronic

The initial reaction to the proposition that electric currents flow in biological organisms was understandably negative. It seemed like proposing that paraffin conducts electricity. Indeed, in the early investigations of this subject, the common view seemed to be justified, for proteins that were crystallized out from the dissolved form and thoroughly dried show an extremely high resistivity.

However, some evidence of a significant electrical conductance in biomaterials was already available in the 1960s. For example, significant conductance was found (Digby, 1965) in crustaceans. Indirect support also came from mechanisms involving electron flow which seemed necessary to explain phenomena in photosynthesis, in enzyme reactivity, and in the energy-producing activities in mitochondria.

---

[4]In corrosion, electrons pass within the metal between cathodic and anodic sites, but the electronic conduction in metals is so high that any potential differences due to this passage can be regarded as negligible. This may not be the case for the application of similar ideas to biological membranes.

A breakthrough came in 1969 with the work of Rosenberg and Postow, who realized that proteins in the body are far from dry, and a nearer approximation to the *in vitro* situation would be to find out the electronic conductivity of *wet* proteins. A difficulty arose here, however, because since the protein is wet, the conductance measured may be partly ionic.

Rosenberg and Postow managed to separate out the ionic component of conductivity. They found that the energy of activation for the electronic part of the conductance went linearly downward as the water content of the protein increased. It appeared as though water, or its ionized constituents, acted like a doping agent in increasing the electrical conductance of intrinsic semiconductors (compare the 1941 suggestion of Albert Szent-Gyorgyi[5] that proteins should be regarded as semiconductors). Thus, an increase of $\sim 10^8$ times in conductance can arise in this way (Fig. 14.11).

---

[5]Albert Szent-Gyorgyi, who continued to work in the laboratory well into his 80s, is ranked as one of the more significant biochemists of the century. The earlier part of his long research life was spent in Hungary, where he became the first person to find that the mysterious factor in limes that eliminated scurvy, vitamin C, was, in fact
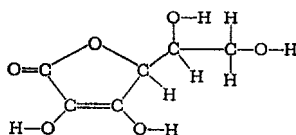
In the 1930s Szent-Gyorgyi was ardent in his opposition to Nazi philosophies. So famous had his name already become that this aroused the personal hostility of Adolf Hitler, who sought to have him arrested, thus triggering his escape and emigration to the United States.

In this country he continued to work far past a normal retirement age, and spurred on by the influx of abundant research money from an active foundation, and marriage to a woman considerably his junior, Szent Gyorgyi turned his independent mind to the problems of cancer. His approach (which he called the electronic theory of cancer) did not deny that DNA and specific genes could determine the likelihood that a given person would be predisposed to cancer. However, he insisted on an electrochemical mechanism for the failure of the normal mechanism of biological cell function. Szent-Gyorgyi pointed to the O_2 reduction reaction ($O_2 + 4H^+ + 4e_0^- \rightarrow 2H_2O$) as the vital step in the functioning of the energy-producing cells. Carcinogens acted as a poison in this reaction.

However, in respect to the mechanism for the reduction of O_2 in biosystems, Szent-Gyorgyi suggested that methyl glyoxal,

$$H-C=0$$
$$|$$
$$C=0$$
$$|$$
$$CH_3$$
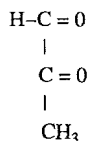
rather than O_2, accepts electrons from enzyme surfaces and thereafter reduces O_2 homogeneously. This kind of biochemical approach to cancer predated the recognition of DNA, which redirected then attention from biochemistry to genetics.

Szent-Gyorgyi is recognized for his seminal suggestion of 1941 concerning the electronic properties of proteins. He supported the concept of hydrogen as the principal storage medium for solar energy (Chapter 15). Thus, he pointed to biomass (formed by means of photosynthesis) as the storer of the sun's energy, by means of the photosynthetic reaction $n CO_2 + n H_2 O \xrightarrow{h\nu} (CH_2 O)_n + n O_2$.
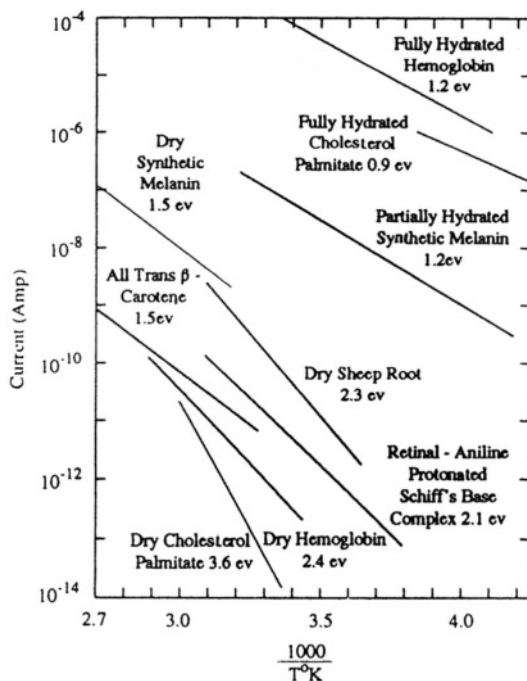
Fig. 14.11. The log of the current plotted against
1/T (K) for a variety of biological substances. The
activation energies are calculated from the slopes
of the lines. (Reprinted with permission from B.
Rosenbergh and E. Postow, *Ann. N.Y. Acad. Sci.*
158:61, 1969, p. 162, Fig. 1.)

At this point in the account of electrical conductance in biological organisms, it
is important to stress the usefulness of examining semiconduction by means of the
Hall effect. Thus, if an electric current flows through a body in the x direction, and a
magnetic field is applied normal to the body in the z direction, an extra electric field
is caused to exist in the y direction. From measurements made, then, on the effects of
applying this magnetic field to a current flowing through a body, it is possible to
determine the charge carrier concentration in the body. Moreover, the sign of the
charge carriers, whether electrons (negative) or holes, (positive) can be determined.
Knowing the conductance, the applied field strength and the concentration and sign
of the charge carriers, it is possible to calculate their mobility. Values range between
$0.01 \ cm^2/V \ s$ to $10 \ cm^2/V \ s$ at room temperature. The carrier concentration varies over
a range of 10^8–$10^{18} \ cc^{-1}$. The biological materials measured (a comprehensive data
set was gathered for the first time by Gutmann Lyons and Keyzer in 1983) show a

wide variety of behaviors as to conductance in which either electrons or holes predominate.

How does electronic conductance occur in biological materials? There are ordered structures in proteins, repetitive groups that allow elements of a band structure to develop. In this respect, the basic mechanism of conductance in an impurity-containing protein is similar to that of a doped semiconductor such as silicon (Bearden and Goldstein, 1986). However, complications occur in respect to the distance between sites from which electrons originate and those in which they may be received. Basically, electrons tend to transfer through energy barriers within the structure by quantum mechanical tunneling. However, as a rough rule, the maximum jump length for the tunneling electron is about 20 Å. If there is no receiver state for an electron within that distance, the probability of a successful tunnel transfer will be so low that the dominant mechanism of transport may change from that of tunneling to that of electron hopping between neighboring atoms. Systems that allow tunneling will generally show a much higher value of charge-carrier mobility than those in which atom-to-atom electron hopping is the preferred mode. Another factor is crystallographic and concerns the shape and size of ordered areas in the protein in which there are repetitive groups, giving rise to band formation, which then enhances the concentration of electrons able to travel under a field gradient.

14.3.2. Protonic

There are situations in biological structures in which the charge carrier is neither an electron nor a hole, but a proton (Nagle and Morowitz, 1978; Pethig, 1998). Because quantum mechanical tunneling is a frequent mode of transport with electrons, it may be asked if it is likely that protons could also tunnel? Thus, the Gamow factor for the probability of tunneling is given by Eq. (9.14):

$$P_T \alpha \; e^{-\frac{4\pi\ell}{h}\sqrt{2mE}}$$

where m is the mass of the tunneling particle, ℓ is the transfer distance, and E is the energy of the entity tunneling. The proton is ~1840 times heavier than the electron. Assuming that the distance of transfer for the proton is about one-tenth that for the electron, and the E value is the same, the tunneling probability for individual protons will be about e^{-4} that for electrons. However, in proteins, even if they are wet, so that the activation energy for electrical conductance is only ~0.5 eV, the probability of activating an electron to the conductance band will be as little as about 10^{-7} at room temperature. It may be that (particularly in proteins) the proton concentration exceeds that of conductance-band electrons so that conductance via a proton tunneling mechanism becomes likely. Detailed investigation suggests that the rate-determining step is usually some conformal change rather than the tunneling step. This would be similar to the situation of protons conducting in aqueous solutions (Section 4.11.4), where the

Direction of proton flow
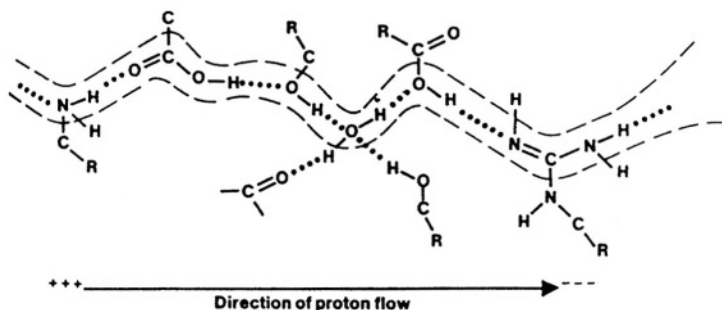
Fig. 14.12. A series of H bonds (···) serving as a proton-conducting pathway within a cytoplasmic or intramembrane protein. (Reprinted with permission from J. F. Nagle and H. J. Morowitz, *Proc. Nat. Acad. Sci. U.S.A.* **78**, 298, 1978, p. 299, Fig. 3.)

rotation of a water molecule under the field of the approaching proton is rate determining, not the rate of proton tunneling (although tunneling occurs after the confirmed change has presented a suitable site for acceptance of the tunneled proton). Similar conformational changes are necessary in a protein structure, in which electron and proton transfer are the important steps in electrical conductance. A possible proton-conducting pathway within a cytoplasmic membrane is shown in Fig. 14.12.

Specific conductance in biological organisms covers the enormous range of 10^{15}. It is possible to understand why the range is so large. The specific conductance will depend on the proteins' structure. Is it suitable for band formation, as it seems to be in many proteins? However, as amply demonstrated, it will also depend on the degree and type of doping, not the intentional addition of trace additives as in the organic semiconductors, but the presence of other entities (water, extra organic substances, fragments, etc.) in the structure. Finally, protons may add to the transport because (although the probability of tunneling of the individual proton is much smaller than that of an electron) the *number* of available protons per unit volume may be far greater.

14.4. THE ELECTROCHEMICAL MECHANISMS OF THE NERVOUS SYSTEM: AN UNFINISHED SECTION

14.4.1. General

Seeing a red light ahead, a driver pauses, then presses the brake with a foot. How long is this pause? 0.5 s? Supposing the time of comprehension to be < 0.1 s, this means that the message to brake traveled from brain to foot muscles at about 4 m s^{-1}. Helmholtz (1989) was the first to publish a scientific measurement of the velocity at

which an electric current passed along the nervous system. His result was "a few meters per second."[6]

At first, it seemed rather simple. Mammals are wired machines, the brain telegraphs the muscles what to do. However, many unsolved problems concerning the passage of information along nerves remained, and still remain. That the signals pass *electrically* is clear enough. But they do not pass as electrons through a metallic conductor, for there are no metallic conductors in the nervous systems of mammals, and the rate of passage is 10^8 times slower than that of electrons through a wire. Electrical activities are an integrated part of the activity of the brain, but there the amplitude of the encephalographic oscillations is in the microvolt region, whereas tens of millivolts are needed to trigger passage of a pulse through the nervous system. It was Bernstein (1902) who first focused attention on the alleged importance of the ratios of the Na^+ and K^+ concentrations of the intra- and extracellular fluids of the nerve axon and related them via a Nernst-type equation to the electrical potential measured across it. Thus the theoretical approach to the passage of electricity through nerves became electrochemical.

Progress in the study of this was very slow until the late 1940s. But then there occurred a grand success which later turned out to be an albatross in disguise. In 1963, Hodgkin and Huxley, two of a small group that had published seminal experiments on the current–potential relation across the membrane of the nerve sheath during the triggering of a nerve impulse received the Nobel prize for their work conducted in 1952. Attached to the original and elegant experiments they described was a phenomenological theory of the surprising results reported. This purported to be a theoretical interpretation of the variation of potential across the membrane containing the intracellular fluid when the passage of a current through a nerve was triggered. The "theory" consisted simply of phenomenological statements of what had been found experimentally. However, it was expressed in mathematical form and was suitably impressive. Everything fitted. There then occurred the phenomenon of Nobel Inflation and Crystallization. The work (identified mainly with the theory) had received the ultimate accolade. It was not then something to be questioned. As the years went on, the theory served its primary users (the electrophysiologists, which are not a small group) well, and thus became beatified: the Hodgkin–Huxley (H-H) theory of the nervous system was an example of bioelectrochemistry glittering with virtue, a work truly to be compared in status with the best-known piece of all electrochemistry, the Debye–Hückel theory of ionic solutions.

The lifetime of good theories in physical chemistry is 50–100 years. The H-H theory is now seen to contain a distinct flaw and a piece of pretense. The equation on

---

[6]Modern measurements distinguish two groups of velocities. Nerves in most mammals have a myelin sheath interrupted by sections in which the nerve axon is in direct contact with the entire cellular fluid ("nodes"). The electrical impulses in the nodes travel at $< 0.01 \text{ ms}^{-1}$ while those in the myelinated sections travel at $\sim 100 \text{ ms}^{-1}$. Evidently Nernst measured a net velocity that would include the effect of interruption of the message by the axons.

which the fit of the theoretical model with experiments was based turned out (Jahn, 1962) to be unable to represent changes in potential resulting from artificially altering the Na^+/K^+ ionic ratios across the membrane. When the glitter had dimmed sufficiently for a critical look, it was seen that the theory is untestable, for it contains a series of coefficients (the ionic permeabilities) P_i that the authors chose to change at will, so that the equation would represent the undisputed facts they had discovered! The electrophysiologists still use the H-H theory—they have no other—but the truth is that this very important piece of electrochemistry will enter the new century as an emperor denuded of his clothes.

14.4.2. Facts

A schematic of a nerve cell is shown in Fig. 14.13. The transfer of information in the body occurs by means of action potentials along the axon.[7] The outside of the latter is sometimes in contact with the extracellular fluid, but for the most part it is sheathed by a substance called *myelin.* The axon is the longest part of the length of the nerve (which eventually ends in a muscle to which its message is given). However, the axon is attached to a nucleus and starlike projections called *dendrites.* Each section of a nerve is connected to the next via a synapse, a space of ~20 nm across which a chemical (e.g., acetylcholine) carries the signal brought by the movement of the action potential along the section.

The structure of a nerve is not simple. In the following account, the stress is upon a single aspect of the mechanism of the action of a nerve, the origin of the spike potential in sections of the nerve called *nodes* in which the axon is in contact on the "outside" with the extracellular fluids. The relevant properties of a nerve cell free of a myelin sheath can be seen in Table 14.1.

The classical experimental work in this area is that of Cole (1949) and in particular that of Hodgkin and Huxley (1952). Their work concentrated upon the potential difference generated across the wall of the axon, which they regarded as a partly permeable membrane. A schematic of their apparatus is shown in Fig. 14.14. This equipment allowed current pulses to be sent across any section of the wall of an axon that was regarded as a membrane.

The key point is that when these current pulses cross the membrane, the potential difference between the solutions inside and outside of the membrane changes significantly. It is this change in potential (and the development of the "spike potential") that is the principal object of discussion in the following paragraphs because the movement of the spike potential along the axon is regarded as the essential act in the transmission of information in the nervous system.

---

[7]Nerves are made up of neurons, cells characterized by long sections (axons) specialized for conducting impulses. Both the cell body (which contains the cell nucleus) from which the axon originates and the end of the axon have many contacts with other cells.
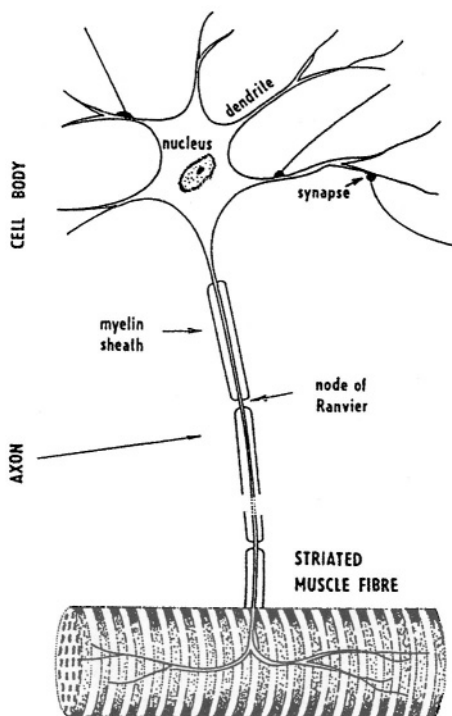
Fig. 14.13. A scheme of a frog motor nerve cell. (Reprinted from J. Koryta, *Ions, Electrodes and Membranes*, Fig. 92, Copyright © J. Wiley & Sons, Ltd. 1991. Reproduced with permission of J. Wiley & Sons, Ltd.)

At the beginning of a series of measurements, the membrane (the wall of the axon) is "at rest," i.e., in its natural undisturbed state. It is found that this is about –70 mV, i.e., the solution on the *inside* of the axon is negative.[8] As shown in Table 14.1, the K^+ concentration is large on the inside and smaller on the outside, while the Na^+ is small on the inside and larger on the outside.

Thus, the phenomena that Cole and Hodgkin and Huxley observed pertained to the behavior of the potential difference between the two solutions, respectively, in contact with the outside and inside of the axon. Because they could not conceive why the unexpected changes in potential should arise *in* the solution (they neglected the

---

[8]A convention is used in the terminology of the potential changes. Hyperpolarization means that $V_{inside} - V_{outside}$ has grown more negative. Depolarization means that $V_{inside} - V_{outside}$ has grown less negative.

TABLE 14.1
Properties of Typical Nerve Cells

Na$^+$ concentration	$c_i = 50$, $c_o = 500$ mM
K$^+$ concentration	$c_i = 400$, $c_o = 10$ mM
Cl$^-$ concentration	$c_i \approx 100$, $c_o \approx 500$ mM
Thiness	5 nm
Resistance	$10^3 \ \Omega$ cm^{-2}
Capacitance	$1 \ \mu$F cm^{-2}
Potential	60–70 mV, – inside, + outside
Rate of propagation of a potential pulse	Order of m s^{-1}
Hyperpolarization	$V_{\text{inside}} - V_{\text{outside}}$ more negative
Depolarization	$V_{\text{inside}} - V_{\text{outside}}$ less negative

Source: Reprinted from J. O'M. Bockris and S. U. M. Kahn, *Surface Electrochemistry*, Plenum, 1993, p. 71.
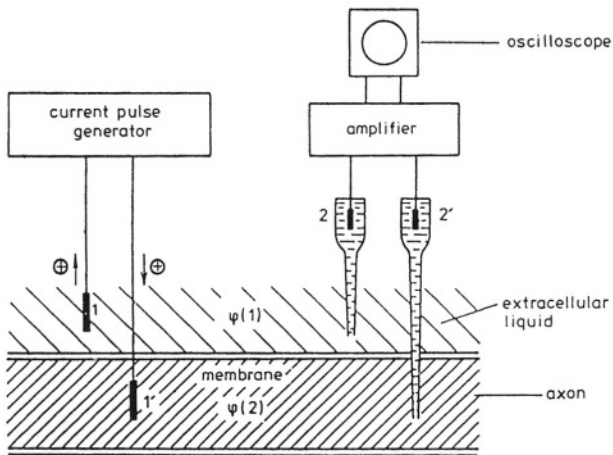
Fig. 14.14. Unit for potential transient measurement during excitation of a squid axon by current pulses from electrodes 1 and 1'; 2 and 2' are micropipettes. (Reprinted from A. L. Hodgkin and A. F. Huxley, *J. Physiol.* **116**:497, 1952. Reprinted from J. Koryta, *Ions, Electrodes and Membranes*, Fig. 93. Copyright © J. Wiley & Sons, Ltd. 1991. Reproduced with permission of J. Wiley & Sons, Ltd.)

possibilities of changes at the interfaces), Hodgkin and Huxley supposed that the changes originated "across the membrane." The changes in the current applied across the membrane and the resultant potential changes are shown in Figs. 14.15 and 14.16.

When several small negative current pulses are passed across the membrane (Fig. 14.15), the potential changes, becoming more negative than the negative rest potential; it becomes hyperpolarized. Then, when the direction of the current pulse changes to positive, the potential grows less negative, i.e., more positive ("depolarization").

However, when the current increases sufficiently in the positive direction, there arises a threshold value at which the potential changes abnormally; it increases to a greater degree than expected, and this larger than expected potential peak is called the *spike potential* (or *action potential*). It is this spike potential, triggered into existence in the H-H experiments by the passage of a sufficiently high positive current across the membrane wall, which then takes off *along* the axon until it reaches the relevant muscle. The movement of the spike potential is the way the message is delivered. The current–time plot at constant potential (Fig. 14.17) greatly resembles a potentiostatic transient at a metal/solution interface (see Section 7.5.10).

14.4.3. The Rise and Fall of the Theory of the Spike Potential

In order to understand the model proposed by H-H (and its continued acceptance by electrophysiologists), it is necessary to remind the reader of changes that have occurred in the electrochemical scene since 1952. With the exception of a few university centers in Russia, England, and Germany, the concept of potentials at this time was still dominated by the thinking of Nernst with the Nernst–Planck equation (4.266) for situations involving diffusion under a potential gradient. Electrode *kinetics* as such had hardly been born. The idea of highly conducting doped semiconductors
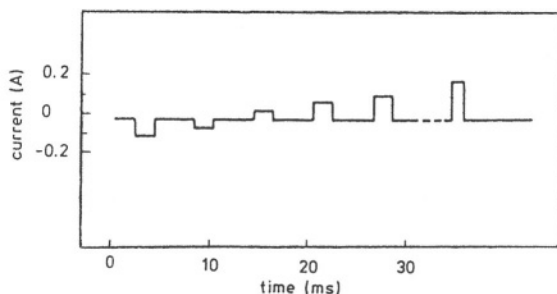
Fig. 14.15. Stimulating current in membrane potential experiment Hodgkin and Huxley (Reprinted from J. Koryta, *Ions, Electrodes and Membranes*, p. 174, Copyright © J. Wiley & Sons, Ltd. 1991. Reproduced with permission of J. Wiley & Sons, Ltd.).
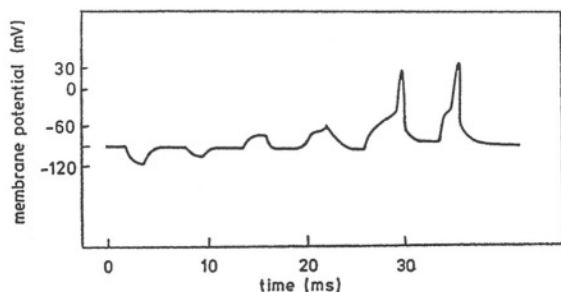
Fig. 14.16. Resulting potential in experiment showing spike in potential that occurs as a result of a critically valued current. (Reprinted from J. Koryta, *Ions, Electrodes and Membranes*, Fig. 94. Copyright © J. Wiley & Sons, Ltd. 1991. Reproduced with permission of J. Wiley & Sons, Ltd.)

was not yet in view. Charge transfer across insulators in contact with solutions containing redox ions (Kallman and Pope, 1960) would have been regarded as irrational.

Thus, in 1952, the equation used to express membrane potentials could be expressed (cf. Eq. 14.5) as:

$$E = \frac{RT}{F} \ln \frac{P_i c_{\text{in}} + P_j c_{j,\text{in}} + \ldots}{P_i c_{i,\text{out}} + P_j c_{j,\text{out}} + \ldots} \tag{14.7}$$
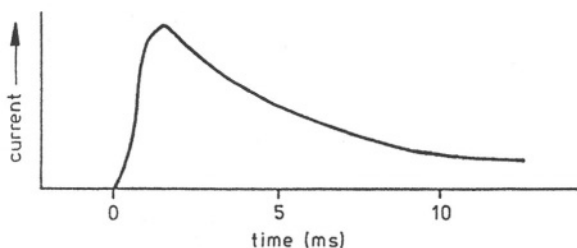
Fig. 14.17. Time dependence of the membrane current. Since the potassium channel is blocked, the current corresponds to sodium transport. (Reprinted from J. Koryta, *Ions, Electrodes and Membranes*, Fig. 95. Copyright © J. Wiley & Sons, Ltd. 1991. Reproduced with permission of J. Wiley & Sons, Ltd.)

Here the permeability coefficients, P, were varied by H-H so that the value of E would fit their observed results. Thus, P's could change with circumstance, with current passed across the membrane, and with time. Insofar as the P's changed, there would be a transfer of K^+ from inside to outside and of Na^+ from outside to inside the membrane. By allowing themselves to adjust the P values to fit their observations, without as yet any indication that the values suggested "opening of the gates," H-H were, of course, easily able to replicate any results they found. Their theory was a sophisticated example—presented with confidence and weight—of the use of the adjustable coefficient. It was followed by the assertion that since the P's had to change in this and that way to fit the results via Eq. (14.7), various movements of Na^+ and K^+ between the inner and outer fluids in contact with the membranes to correspond to the permeation changes must be occurring. Opening and closing of channels ("gates") to allow or disallow entry and exit of Na^+ and K^+, respectively, was then simply assumed as a part of the natural happening, with the expectation that in succeeding decades evidence for such gates might be found.

All this held up well for 10 years or so—and still holds sway among electrophysiologists. However, the Goldman-related Eq. (14.7) was tested directly by Jahn (1962). The axons in a giant squid are particularly large, 1 mm in diameter, and large enough for direct experiments in which micropipettes could be used to artificially change the K^+/Na^+ ratios both inside and outside the axon. Only the solution concentrations were changed so that there was no excuse for any change in the P's. Although Jahn changed the concentrations enough for a change of 50 mV to occur according to Eq. (14.7), the actual change was only a few more millivolts.

However, the final blow to the applicability of the Goldman equation and its use in the H-H theory came when experiments were made with the transport across the membrane of radioactive ^{22}Na. It was found that when the threshold value was reached and a spike triggered, there was indeed *some* movement of ^{22}Na from the outer to the inner fluid, but the amount was about 500 ions cm^{-2} of the axon and the change in potential that this would bring about in the extracellular fluid was tiny, so that when substituted back into the Goldman equation, it could in no way account for the changes of 50 mV or more observed in formation of the spike potential.

Such disproving experiments are sufficient to end acceptance of the theory in physical electrochemistry. New interpretations must be used to explain the stimulating and interesting experimental results that the 1952 experiment revealed, for a theory of the mechanism of the nervous system is a primary problem in science.[9]

Have the intriguing findings of H-H attracted bioelectrochemists away from the goal? What is that goal? It is to find a scientifically tenable model for the passage of electricity *along* nerves. H-H concentrated on something different than that: why an

---

[9]Not only in science but in philosophy, too. Thus, a mechanistic theory of the *transfer* of information still begs the question of who decides to decide. Who (what?) creates the 30-mV pulse needed to trigger a spike and start a message flow, although encephalographic oscillations in the brain have a range of microvolts?

inexplicably large potential *across* the axon's wall could be triggered if the values of an artificially induced current *across* the membrane wall reached a critical value. Is there an analogy between their concentration on this finding and that of the lady who lost her diamond just outside the theater entrance, but looked for it up the road because there she had a street light?

If it is argued that an explanation of the genesis of the spike potential that travels down the nerve must be the answer to nerve conduction, then the reader must be reminded that it cannot be more than a part of that answer. What of the sections of the nerve where there is no contact with any extracellular fluid? What about the synapses, where, for 20 nm, the charge is carried by large, lumbering molecules? Would not the rate-determining step for the transfer of electricity in nerves be *there* and not in the small sections that are nodes—a contact with the outer cellular solution?

Very complex and sophisticated biochemistry has been devoted to finding out the minutiae of the channels through which it is alleged that ions migrate (diffuse?). Among the findings are movements against the electrochemical potential gradient ("active transport") and here surprisingly mobile and intricate (purposeful?) movements of enzymes are hypothesized: K^+ ATPase is supposed to seize a K^+ and hurl it in a direction against that indicated by the Nernst–Planck law.

One of the more substantive attempts to rationalize the alleged openings and closings of channels in the membrane has been given by Blank. Blank's idea is expressed in terms of oligomers, preproteins that consist of stacks of amino acids (e.g., tetramers) (Fig. 14.18). The channels consist of proteins that can undergo conformal changes, depending on pH (and temperature), but particularly due to charge (Fig. 14.19). In Blank's model, the migration of charge from the surface (channel closed) to the interior (channel open) is the rate-determining event in nerve conduction. Thus, some rationale for the opening and closing of channels may be given.

Finally, in attempting to explain the forward movement in nerve conduction, one may formulate two sets of questions:

1. A potential spike of about 100 mV travels down the nerve. How does such a large potential get triggered in the brain? The H-H method of creating this action potential in the laboratory by passing a current laterally across the axon wall seems irrelevant to the *in vivo* situation.
2. By what molecular mechanism does the action potential move?

As to the forward movement, this is described in the literature on the basis of the H-H theory as follows:

The action potential is self-propagating because at the peak of the action potential, when the inside of the membrane at the active region is comparatively positive, positively charged ions move from this region to adjacent areas inside the axon, which are still comparatively negative. As a consequence, the adjacent
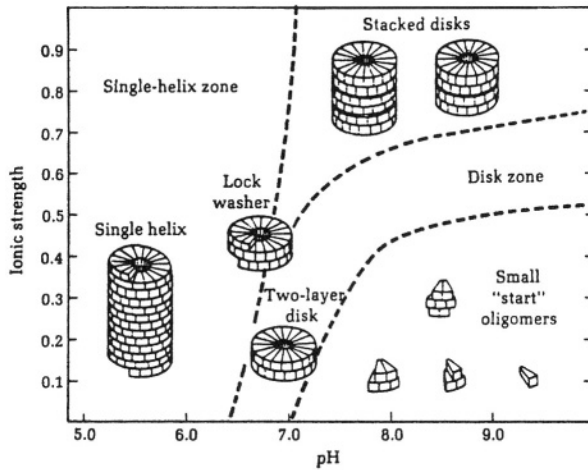
Fig. 14.18. Diagram due to A. Klug showing the stability of aggregates of tobacco mosaic virus (TMV) protein over a range of ionic strength (ordinate) and pH (abscissa). The boundaries are approximately correct for a protein concentration of 5 mg/ml at 20 °C. The isoelectric point of the TMV protein is 4.7, so the diagram shows greater aggregation with decreasing negative charge or increasing ionic strength. (Reprinted from M. Blank and L. Soo, *Bioelectrochem., Bioenerg.* **24**: 51, copyright 1990, with permission of Elsevier Science.)
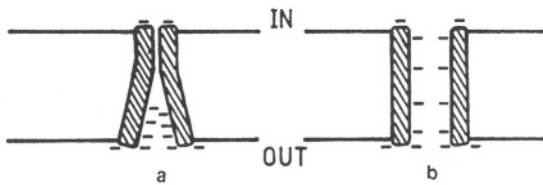
Fig. 14.19. A diagram of a voltage-sensitive oligomeric channel within the membrane matrix. (a) In the resting state, the asymmetric distribution of charge density causes the outer region to be dissociated while the inner region is associated. (b) Upon depolarization, the charge density shifts, causing the inner region to dissociate and the channel to open. (Reprinted from M. Blank and L. Soo, *Bioelectrochem., Bioenerg.* **24**:51, copyright 1990, with permission of Elsevier Science.)

area, in turn, becomes depolarized—that is, less negative. When it becomes slightly depolarized, its permeability to Na^+ increases. Na^+ ions rush in and create a new action potential, which, in turn, depolarizes the next adjacent area of the membrane. Thus, the nerve impulse travels along the axon.

The segment of the axon behind the nerve impulse has a brief refractory period during which it cannot be reexcited and which keeps the action potential from going backwards. Because of this renewal process, repeating itself along the length of the membrane, an axon, which would be a very poor conductor of an ordinary electric current, is capable of transmitting a nerve impulse over a considerable distance with absolutely undiminished strength (Curtis, 1979).

Such an explanation collides with the fact that only a few ions transfer in the formation of the spike potential. Or, "The spike radically changes the electric field in the surrounding of the channel and causes depolarization also in neighboring channels, thus making possible the transfer of the impulse along the axon" (Koryta, 1991). Such explanations obfuscate with words. They do not give a model that can be grasped at a molecular level.

Thus, in regard to this important part of electrophysiology, the mechanistic details are still fuzzy. What of the conductivity of the axon material? Is it ionic? Electronic? Can there be some kind of Grotthus mechanism such as that illustrated in Figs. 4.121 and 4.123 in Vol. 1?

There is still much to be learned about the changes in potential across the nerve axon that are associated with conduction of electricity along nerves. The first step is to question the relevance of concepts connected with the lateral motion of alkali metal ions *across* the axon wall (for it happens to a minute extent) and concentrate on the origin of the spike *in vivo*. In parallel experiments, the mechanism of interfacial charge transfer involving each side of the axon with a saline-type solution containing appropriate organic constituents should be studied. Such a study would seem to offer much opportunity for modern methods of surface examination, in particular radiotracer measurement aided by sectioning of the axon and use of XPS, atomic force microscopy, and other modern methods of examining the biological/solution interface (Roscoe, 1996).

The demise of the famous Hodgkin–Huxley theory of nerve conductance brings to mind other Nobel prizes in electrochemically related areas. In 1959 Heyrovsky was recognized for a new analytical method, and this polarography has been the origin of many modern methods of electroanalysis. The award for Nobel Prize to Mitchell in 1978 (for a "chemiosmotic" model of membrane function) and metabolism seems to have been based on a lack of awareness of a simpler, clearer (prior) model by Williams for interpreting the same functions. The award to Marcus in 1992 for the theory of redox reactions (1956) seems to have lacked awareness of an earlier publication by Weiss that described similar ideas.

14.5. INTERFACIAL ELECTRON TRANSFER IN BIOLOGICAL SYSTEMS

14.5.1. Introduction

What is developing at the frontiers of concepts of reactivity in bioelectrochemistry is a view of biochemical reactions in which interfacial electron transfer (often at enzyme/solution interfaces) plays an important, sometimes rate-determining, role (Bockris, 1969). The theory of such electrochemical reactions is more complex man that for the metal/solution situation and two reasons for this are worth identifying. (1) On the electrode side (e.g., the solid surfaces of a protein), the surface is more complex and its electronic properties are less well understood than those of a metal. (2) The "ions" in solution, to and from which electrons are exchanged with the biosurface, are usually very large, and have equivalent radii 10–100 times greater than those of the ions and molecules of normal electrochemistry. On the other hand, it seems likely that interfacial electron transfer is often a rate-determining step in biological processes, and its study introduces what Szent-Gyorgyi called "the electron-level theory of life."

Simplification is necessary for understanding most real processes. In this case the description begins with material on the adsorption of biomolecules on metals; then we discuss the active field that has developed in a study of electron transfer from modified metal electrodes to proteins dissolved in solution; finally we describe the as-yet less well developed study of charge transfer from proteins to simple redox ions in solution. The *real* field of course is the kinetics and mechanism of electron transfer from proteins to biomolecules, but this area of experimental research is as yet a bridge too far.

14.5.2. Adsorption of Proteins onto Metals from Solution

Of the dozen or so methods that have been used to examine the adsorption of biomolecules onto metals from solution, those most suited seem to be (1) cyclic voltrammetry, (2) ellipsometry, (3) Fourier transform infrared spectroscopy, and (4) atomic force microscopy.

Because it is the model protein, cytochrome *c* is the most examined protein[10] in respect to surface studies. It absorbs apparently without dissociation to the extent of a monolayer on fluorine-doped SnO_2 electrodes (Willit and Bowden, 1990).

Some significant experiments on the adsorption of biomolecules on metals were carried out by Szucs et al. (1991). They used ellipsometry to continuously monitor the surface of a gold electrode as a function of time when the solution of the enzyme glucose oxidase was added to the solution. They also measured the various voltamograms that were seen on the gold electrode during adsorption and subsequent events. Three different stages were distinguished with this technique. In stage one [Fig. 14.20(a)], the enzyme landed on the electrode intact. In stage two, it "fell over," i.e.,

---

[10]This is because its structure has been well established by means of X-ray work. It is relatively small (in comparison with either biomolecules) and takes part in redox processes in mitochondria.
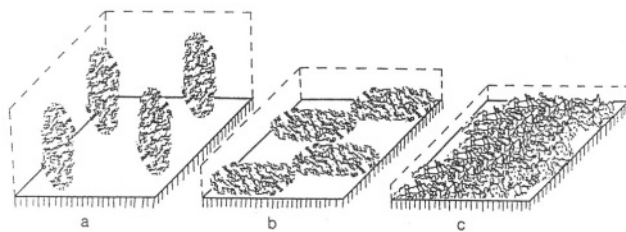
Fig. 14.20. Schematic representation of the adsorption process of glucose oxidase (GO); (a) The potential is much more positive than the potential of zero charge. (b) The potential is close to the pzc. (c) The final stage when the enzyme unfolded. (Reprinted from A. Szucs, G. D. Hitchens and J. O'M. Bockris, *J. Electrochem. Soc.* **136**(12): 3748, Fig. 13, 1989. Reproduced by permission of the Electrochemical Society, Inc.)

the major axis of the enzyme lay parallel to the electrode surface [Fig 14.20(b)]. In the third stage, the enzyme appeared to break apart [Fig. 14.20(c)]. The thickness of the absorbed enzyme, measured ellipsometrically, started at about 140 Å, but in the third stage it was only 25 Å (Fig. 14.21).

At first, a two-electron transfer reaction took place between gold and the glucose oxidase. However, as time (in terms of minutes) went on, the electrodic response (Fig.
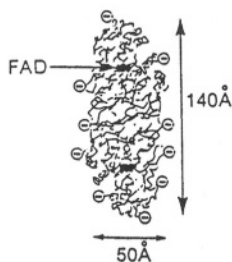
Fig. 14.21. The dissociated enzyme. (Reprinted from A. Szucs, G. D. Hitchens and J. O'M. Bockris, *J. Electrochem. Soc.* **130**: 6(12):3749. Fig. 1, 1989. Reproduced by permission of the Electrochemical Society, Inc.)

14.22) decreased, corresponding to an unfolding of the enzyme, and the electrode reaction corresponded to the reaction of a part of the enzyme, namely, flavin adenine dinucleotide (FAD). Thus, the FAD left the molecule during its breakdown after contacting the surface. The enzyme's unfolding allowed a redox center to absorb on the electrode and react there. This evidence of enzyme dissociation shows that in this case, absorption on a metal has made the enzyme itself no longer active, owing to its breakup on the electrode surface.

Thus, the work of Szucs et al. proved that adsorption of an enzyme on an *unmodified* metal may be disastrous for the enzyme's structure. The absorbed enzyme
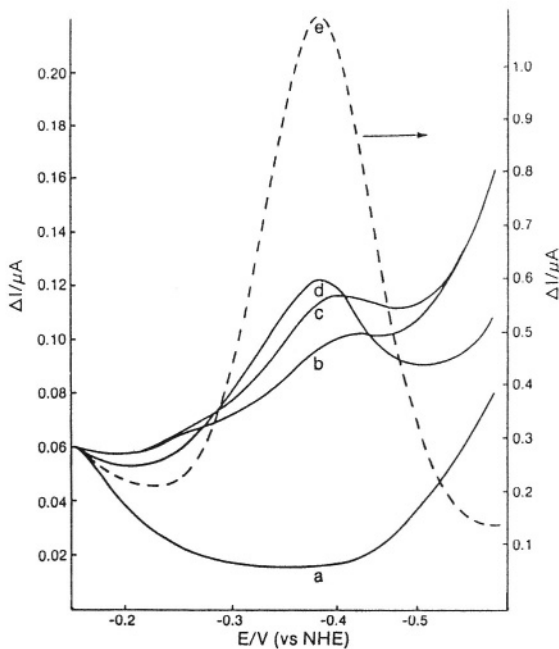
Fig. 14.22. Differential pulse voltammogram of GO and FAD on a gold electrode. a, Background current; b, after 10 min adsorption of the enzyme from 5.3 μM solution; c, after 40 min; d, after 120 min; e, electrochemical response of adsorbed FAD measured in pure buffer after 2 hr adsorption from 10 μM FAD solution and washing the electrode. (Reprinted from A. Szucs, G. D. Hitchens, and J. O'M. Bockris, *Bioelectrochemistry* **21**: 133, copyright 1989. Reproduced with permission of Elsevier Science.)

resembles an aircraft after a crash landing; it can no longer fly, i.e. no longer perform its function of supercatalysis.

What happens to glucose oxidase upon absorption can be understood in terms of the actual structure of an enzyme. Enzymes are relatively complex, and their structures as organic molecules are difficult to draw. However, it is possible to make a representation, although much is lost in the absence of a three-dimensional model. A diagram due to L. Sawyer (1991) of the enzyme β-lactoglobulin is shown in Fig. 14.23.

Ellipsometry is useful in examining the behavior of enzymes on electrodes. Another excellent tool is the atomic force microsope (Section 7.5.18). Eppel (1993) used AFM to examine the absorption of a glycoprotein on some biosurfaces (AFM does not need conduction electrons to function), and was able to measure its elliptical cross section in the absorbed state. It was possible to map changes in the charge density of the protein (which plays a role in thrombus formation in arterial blockage) upon absorption by the change in the height of various parts of the molecule. This is an excellent example of the detail in which absorbed species can be examined.

Much information about protein absorption can be obtained by measuring the electrochemical isotherm, the $\theta - c_{soln}$ relation at constant temperature for a series of fixed potentials. One takes the protein (P) as absorbing fast upon the metal, and transferring n electrons, corresponding to n carbonyl interactions. The dissociation reaction is taken to be rate determining. Then (Roscoe, 1996):
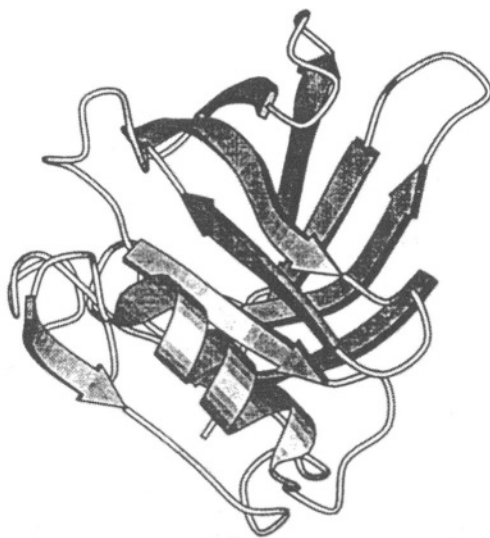
Fig. 14.23. Diagram of β-lactoglobulin. (Reprinted from L. Sawyer, University of Edinburgh, 1996.)

$$P + nM \rightarrow P(M)_n + ne^- \quad \text{(fast)}$$

$$P(M)_n \rightarrow P'(M)_n + nCO_2 \quad \text{rds}$$
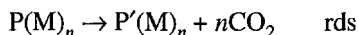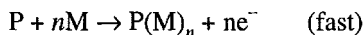
The rate expressions for the forward and reverse reactions, having rate constants k_1 and k_{-1} and free energies of activation $\Delta G_1^{\ddagger}$ and $\Delta G_{-1}^{\ddagger}$, respectively, may be written as:

$$i_1 = nzFk_1(1 - \theta_G)^n C \exp[-\Delta G_1^{\ddagger} - \beta VF + \beta r_1 \theta)/RT] \tag{14.8}$$

$$i_{-1} = nzFk_1(\theta_G) \exp[-(\Delta G_{-1}^{\ddagger} + (1 - \beta)VF - (1 - \beta)r_1\theta)/RT] \tag{14.9}$$

where G refers to the species on the electrode surface, C is the concentration of protein in the bulk, θ is the surface coverage, n is the number of functional groups on the protein and is taken as equal to the number of electrons transferred, β is the symmetry factor, and r_1 allows for variation in the heat of adsorption with coverage.

If step 2 is rate determining, then step 1 can be considered to be in quasi-equilibrium. Hence,

$$\frac{\theta}{(1 - \theta)^n} = KC \exp(-\sigma) \tag{14.10}$$

where

$$K = k_1/k_{-1}$$

$$\sigma = (\Delta G - VF + r_1\theta)/RT \tag{14.11}$$

Thus:

$$\ln(C/\theta) = n[-\ln(1 - \theta)] - \ln K + \sigma \tag{14.12}$$

The value of n may be determined from the slope of a plot of $\ln(C/\theta)$ vs. $-\ln(1 - \theta)$. The slope of 25 (Fig. 14.24) obtained from measurements made using an anodic end potential of 0.4 V corresponds well with the number of carboxyl groups on the β-lactoglobulin A molecule. Similar behavior is found for a number of proteins adsorbing on inert metals.

14.5.3. Electron Transfer from Modified Metals to Dissolved Protein in Solution

Looking at the size and complexity of a protein, as indicated in Sawyer's figure (14.23), and then mentally comparing it with, say, hydrated Fe^{3+}, a typical subject of fundamental electrochemical studies, one might be a bit discouraged as to the hope of
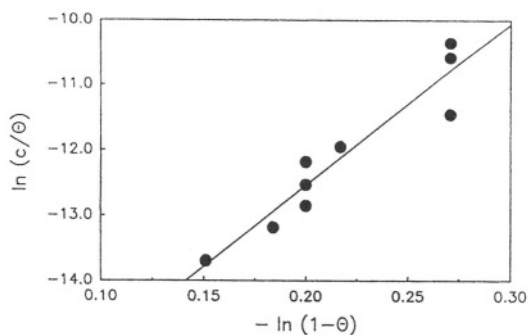
Fig. 14.24. Plot of $\ln(C/\theta)$ vs. $-\ln(1-\theta)$ for β-lactoglobulin in phosphate buffer (pH 7.0, 299 K) with an anodic potential of 0.4 V (vs. SCE). (Reprinted from S. G. Roscoe, K. L. Fuller, and G. Robitaile, *J. Colloid Interface Sci.* **160**:243, 1993, with permission from Academic Press.)

studying electron transfer to biochemicals. The pioneers who made the breakthrough here were Kono and Nakamura at the early date of 1958. (They attempted the reduction of cytochrome c at Pt.) A number of such studies were reported between 1970 and 1980, and it was found (particularly by Hawkridge et al.) that success was more likely at oxide semiconductors (e.g., tin oxide and indium oxide), rather than on metals. They established that the detailed nature of the surface controlled success, rigorous electrode cleaning being necessary.

An important change in this kind of work came in 1981 when it was found (Alberry and A. O. Hill) that modification of single-crystal electrode planes by the preabsorption of a monolayer of a certain kind of organic was helpful in promoting a successful landing of the giant enzymes on electrodes. Thus, if 4-4'-bipyridyl and 4-4'-dithiopyridine were preabsorbed onto graphite and cytochrome c was adsorbed on top of them, it underwent redox reactions at a much faster rate than when the protein adsorbed simply upon a metal or semiconductor substance (Fig. 14.25). The situation can be seen in Fig. 14.26. A number of other functional electrode surfaces used in the study of electron transfer to biomolecules are shown in Fig. 14.27.

The nature of the adsorbed organic—the promoter—its precise structure, chemistry, and interaction with the protein is of vital importance (Taniguchi 1997). This is shown in Fig. 14.28 where atomically flat Au(110) and Au(100) surfaces were covered with several different bipyridyls. The 4,4'-PySSPy gave good response but the 2,2'-PySSPy gave none. It is a matter of the interaction of the absorbing protein with the modifier. If it brings the central metallic heme group close enough to the metal so that quantum mechanical electron tunneling from metal E to the heme group (often Fe^{3+}) can occur, electrons will pass between the electrode and the heme group, i.e., redox reactions will occur.
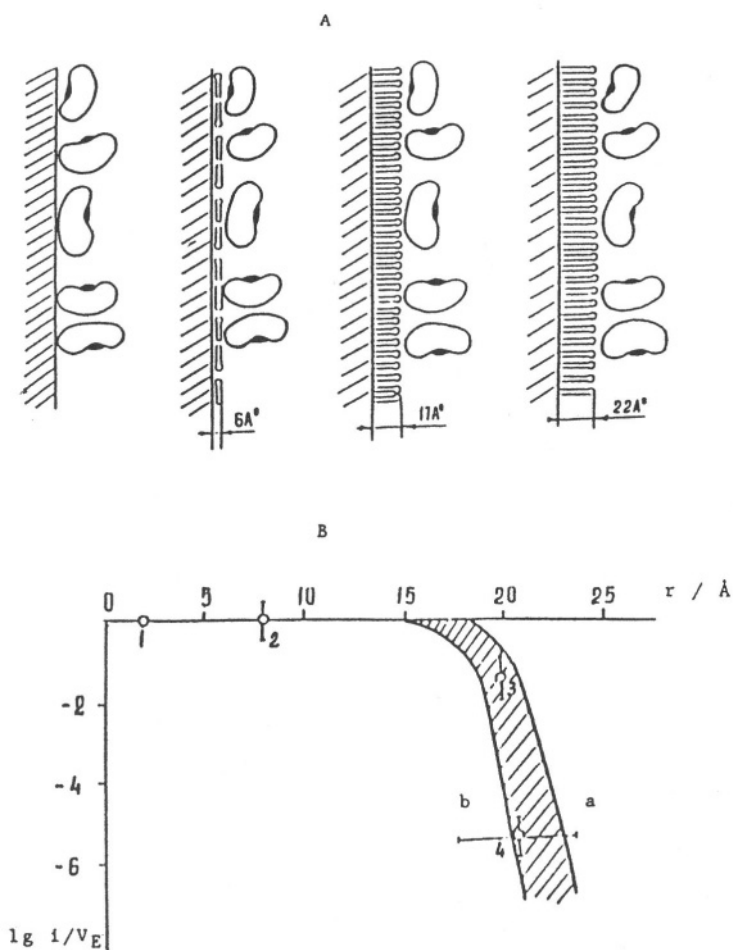
Fig. 14.25. (A). The scheme of enzyme adsorption of the electrode with different lipid interlayers. (B). The relative oxygen reduction rate vs. the distance between the electrode and the enzyme: for lactase adsorbed (1) on soot, (2) on cholesterol, (4) on lecithin. The curves a and b are calculated for barrier heights of 4 and 5 eV, respectively. (Reprinted from J. O'M Bockris, M. Szklonzyk, and Szucs, in *Electropharmacology*, G. M. Eckert, F. Gutmann, and H. Keyzar, eds., Figs. 25, 29, 30, 1990. Reproduced with permission of CRC Press.)
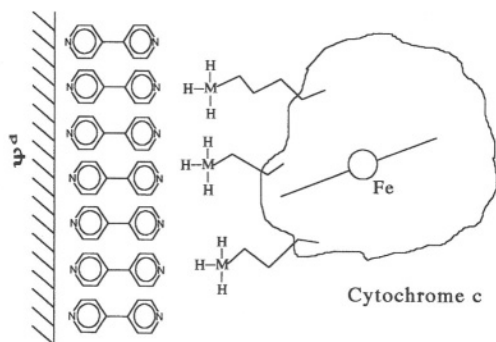
Fig. 14.26. An illustration of the electrode–protein complex envisaged for the idealized limiting case of a planar surface. (Reprinted with permission from F. A. Armstrong, H. A. O. Hill, and N. J. Walton, *Acc. Chem. Res.* **21** p. 408, Fig. 2, copyright 1989, American Chemical Society.)
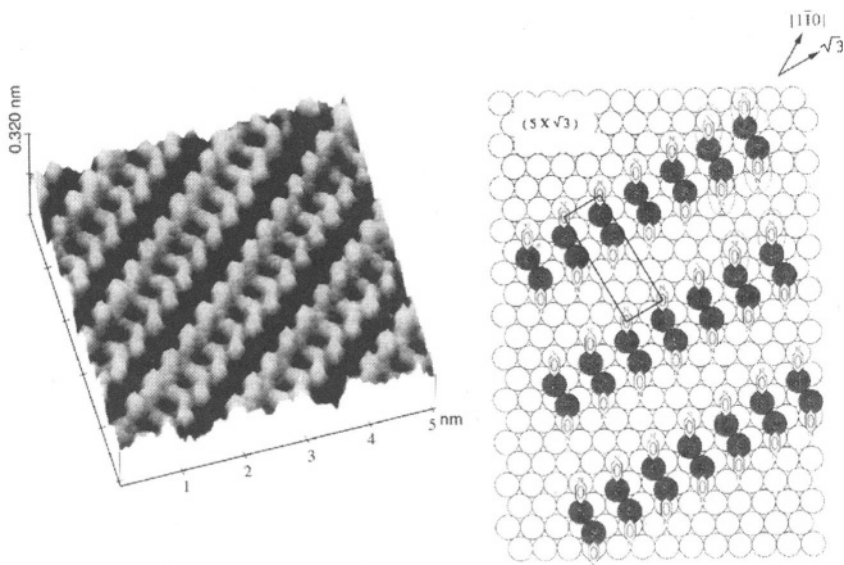
Fig. 14.27. Typical STM image of the 4,4′-PySSPy on an Au(111) surface. (This picture was taken with the aid of Dr. Takahito Sawaguchi of the National Institute of Bioscience and Human Technology. (Reprinted from I. Taniguchi, "Probing Metalloprotiens and Bioelectochemical Systems," *Interfacial* **6**(4): 34–37, Fig. 2, 1997. Reproduced by permission of The Electrochemical Society, Inc.)
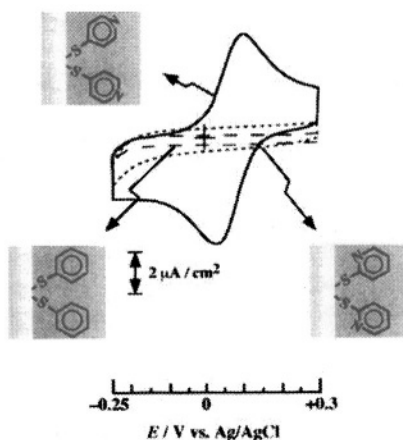
Fig. 14.28. Cyclic voltammograms of 100 µM horse heart cytochrome c at an Au(110) surface modified with 4,4'-PySSPy, 2,2'-PySSPy, and diphenyl disulfide (PhSSPh) in a 0.1 M phosphate buffer solution containing 0.1 M NaClO$_4$ (pH 7) at 25 °C. Scan rate: 50 mV/s. (Reprinted from I. Taniguchi, "Probing Metalloproteins and Bioelectrochemical Systems," *Interfacial* **6**(4):34–37, Fig. 3, 1997. Reproduced by permission of the Electrochemical Society, Inc.)

The future direction of this work is to modify electrode surfaces in such a way that enzymes can absorb upon the modified structures without the breakdown and dissociation monitored by Szucs et al. for glucose oxidase on unmodified gold electrodes. Then it may be possible to use the specificity of enzymes not only in solution, but absorbed on electrodes, just as enzymes absorb on cell surfaces. There may be an analytical function to be developed here; diseases produce specific new molecules that circulate in the blood. It may be possible to develop a number of enzymes, all adsorbed on an electrode, but each reacting specifically only with one disease-specific molecule. A suitably modified electrode, with a connection to micropatches containing the disease-specific enzymes, could transfer electrons back to the underlying metal and hence to specific outer circuits, each one reflecting current gathered from a specific enzyme patch. In this way, a signal would show up upon an external monitor, directly indicating the presence of a specific molecule and disease.

14.5.4. Electron Transfer from Biomaterials to Simple Redox Ions in Solution

Something has been learned so far about electron transfer from a metal to simple ions in solution (normal electrochemistry, Chapter 7), from metals to biomolecules in solution, and from promoter-modified metals to biomolecules in solution. The next step is to examine electron transfer from a biosurface to a simple redox species dissolved in an ambient solution.

Experiments preliminary to those involving actual membranes containing proteins were carried out by Bockris and Schuaib in 1978. These consisted of the isolation of material from photosystem I and photosystem II (the active receptors in photosynthesis) from spinach, a typical green plant, growing by means of the photosynthetic reaction $nCO_2 + {}_nH_2O\, nh\nu \rightarrow (CH_2O)_n + nO_2$. Absorbing these materials individually and separately on platinized Pt and exposing the metal alone and then the metal with either the one or the other of the adsorbed biomaterials showed a significant photoactivity (i.e., electron transfer between the biospecies adsorbed on the electrode and entities in the solution). The photocurrents with one photosystem upon irradiation were anodic. However, when the other photosystem was adsorbed on the electrode and the latter irradiated, there was a cathodic current flow. Such results form the basis of the evidence that the first part of the photosynthesis reaction consists in the photoelectrochemical splitting of water, the provision of H_2 via the cathodic photosystem and oxygen from the other. [11]

Experiments with membranes containing immobilized proteins are more difficult than those in which the proteins are free and dissolved in the solution, later adsorbing temporarily to collect or give electrons to the promoter-modified electrode. In biological cells of living systems, the membranes, some with enzyme layers attached, are extremely thin. It would be difficult to find an experimental arrangement in which such a layer of actual biomaterial could be made into an electrode attached to an outer power source, etc. Because of such difficulties, the examination of electron transfer at the interfaces of biosystems has been a path less traveled.

However, as pointed out earlier, the artificial bilayer membrane, the BLM, can be made to serve as a model for the reality of biomembranes. One forms the BLM itself (Section 14.2) and then introduces into it various entities in order to examine their chemical and electrochemical effects. The appropriate membrane can be assembled by the use of a Langmuir–Blodgett trough in which long lipid molecules (those that make up the bilayer) are floated on the surface of water and then gently pushed together by a plastic slider (Rejou-Michel and Habib, 1986). A sensitive mechanism measures the force of this pushing and then when the force necessary increases suddenly, one knows the molecules have all been pushed into contact and a monolayer formed.

---

[11]The completion of the photosynthetic reaction to form CH_2O and O_2 occurs chemically using CO_2 from the atmosphere and H_2 from the initial photoelectrochemical breakup of water.

What is the point of forming a BLM this way? If one wants to make a bioelectrode of varying constitution, this method has two advantages. It is easy to introduce some protein molecules into the BLM layer and thus make an artificial membrane of more or less any design. Another advantage is more specific to the purpose of forming a bioelectrode. One has to have mechanical strength and also an electronically conducting background so that the intended passage of electrons emitted from the proteins in the BLM, to ions in the double layer in contact with the electrode will not be held up by lengthy passage through a poorly conducting material to which the BLM membrane might be attached. This is achieved by evaporating Au or SnO_2 onto a glass slide, placing this underneath the bilayer formed on the surface of a solution in a trough of the Langmuir–Blodgett apparatus, and lifting it up out of the solution, the Au-covered slide now containing on its surface the BLM membrane with its putatively conducting protein segments. Thus what is formed is a bioelectrode.

There are still many difficulties in such experiments which aim to demonstrate electron exchange between proteins and ions in solution. Ions gradually diffuse into BLM layers. Hence, it is not certain that such a layer protects the ions in solutions from an undesired contact with the underlying electronically conducting coating. Alternatively, there are sometimes some kind of microfissures in such membranes ("pinholes") and one has to be careful that these do not allow a parallel pathway for the redox ions in the solution to come in and get, or give, electrons to the underlying gold (the aim of the experiment is to get them to exchange electrons with the protein additions in the BLM membrane). The various possibilities are seen in Fig. 14.29.

In pioneer experiments of this type carried out by Rejou-Michel and Habib (1982), the two coatings on the Au-coated glass slide were the BLM alone and the BLM with about 12% of gramicidin, a polypeptide. As indicated, a danger with experiments of this type is that current will be passed *through* the BLM membrane. Some results are shown in Figs. 14.30 and 14.31. In the first of these two figures (three layers of BLM containing no gramicidin), the results are similar to those obtained on the Au substrate; the BLM membrane has clearly leaked. When a five-layer BLM is used, the results are similar to those of diffusion-controlled currents; the extra layers have turned the rate control into that of diffusion of the redox ions in solution through the BLM membrane to the underlying gold layer.

In the membranes containing gramicidin, there was no change in diffusion control when a change from three to five layers was made. The Tafel slope changed from 0.16 (no gramicidin) to 0.20 on the membrane containing gramicidin. The i_0 for the latter membrane was twice as great as that on the BLM–Au contact. These results suggest that the gramicidin within the BLM contributes significantly to the electron transfer rate in parallel to reactivity arising from a penetration of ions through the BLM to the underlying metal.

Similar experiments have been carried out by Hawkridge et al. (1989). They used coenzyme Q 10% within a layer of phosphatidylcholine and showed activity of the enzyme as in Fig. 14.32.
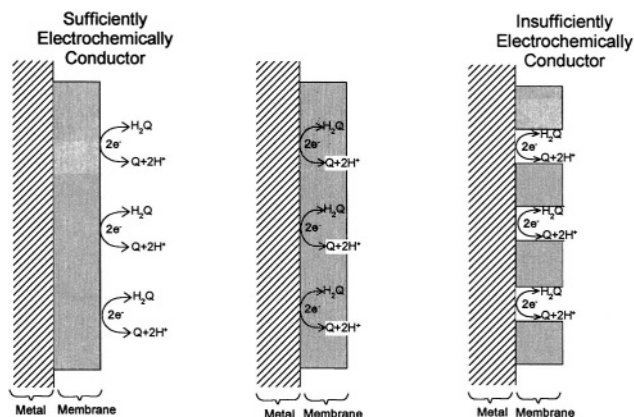
Fig. 14.29. Schematic diagram showing the three possible alternatives for the location of electron transfer. (Reprinted from A. Rejou-Michel, M. A. Habib, and J. O'M. Bockris, "Electron Transfer at Biological Interfaces," in *Electrical Double Layers in Biology,* M. Blank, ed., Fig. 3, p. 171, Plenum, 1986.)

14.5.5. Theoretical Aspects of Electron Transfer from Solid Proteins to Ions in Solution

Two possible mechanisms have been compared for electron transfer from solid-positive to ions in solution (Khan, 1993). In the first, electron tunneling from the metal through the membrane (thickness 70–100 Å) was examined and found to give values many orders of magnitude less than those experimentally observed. In the second, it was assumed that impurity states exist in the protein component of the membrane and that the electron tunnels from state to state. This model is found to be reasonably consistent with experiment.

14.5.6. Conduction and Electron Transfer in Biological Systems: Retrospect and Prospect

Electrochemistry is often thought of as having been born during Galvani's experiments with the frog's legs in 1791. What followed was, however, Volta's experiments with batteries and it was the latter, in dealing with tiny atoms of zinc and silver, that set the stage for the development of electrochemical science. In this chapter on bioelectrochemistry we have gone back to join up with Galvani and explore the possibilities that the potential differences across biological membranes, which have been examined for such a long time, are significant for understanding some aspects of molecular biology.
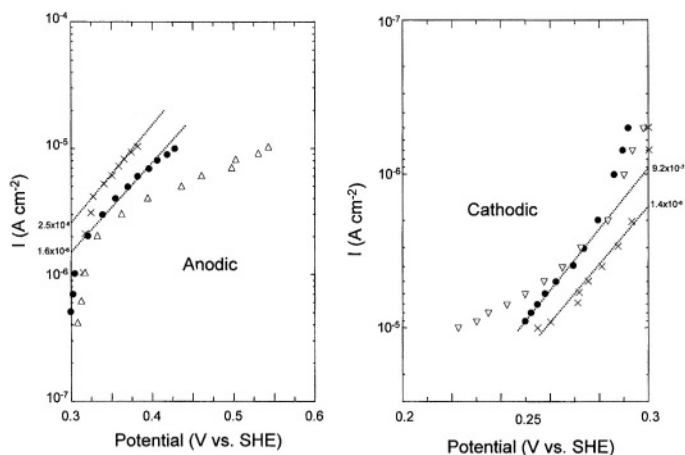
Fig. 14.30. (a) Logarithm of the anodic current density–potential curves of the Q/QH_2 redox couple on ($\times$) gold electrode; ($\bullet$) gold electrode covered by three layers of DPPC, ($\triangle$) gold electrode covered by five layers of DPPC. (b) Logarithm of the cathodic current density–potential curves of the Q/QH_2 redox couple on ($\times$) gold electrode, ($\bullet$) gold electrode covered by three layers of DPPC, (∇) gold electrode covered by five layers of DPPC. (Reprinted from A. Rejou-Michel, M. A. Habib, and J. O'M. Bockris, "Electron Transfer at Biological Interfaces," in *Electrical Double Layers in Biology*, M. Blank, ed., Fig. 8, p. 175, Plenum, 1986.)

However, first we have to point out a couple of differences that distinguish the topics of this chapter from those in the others. (1) The biomolecules we deal with, the proteins and enzymes, are between 10 and 100 times larger than the ions and molecules that we have discussed in the rest of the book. (2) The complexity of biological systems is much greater than the complexity of the systems we have been dealing with, and this makes simple answers more difficult to obtain.

The measurement of membrane potentials has been reported more than any other experiment in bioelectrochernistry. Its history stretches back into the 1900s and the theories that were in place to explain such phenomena are phenomenological. Such theories climaxed in the 1950s with that of Hodgkin and Huxley, who claimed to have explained one of the more prevalent pieces of electrochemistry in nature, the passage of electricity from the brain to the various parts of the body via the nervous system. However, since the enunciation of their theory in 1952, experiments carried out to directly test its assumptions have shown results inconsistent with the theory. Membrane potentials may have a number of origins, but the present view is in the direction of seeing the potential as being the *result* of electrode processes on the surface of the
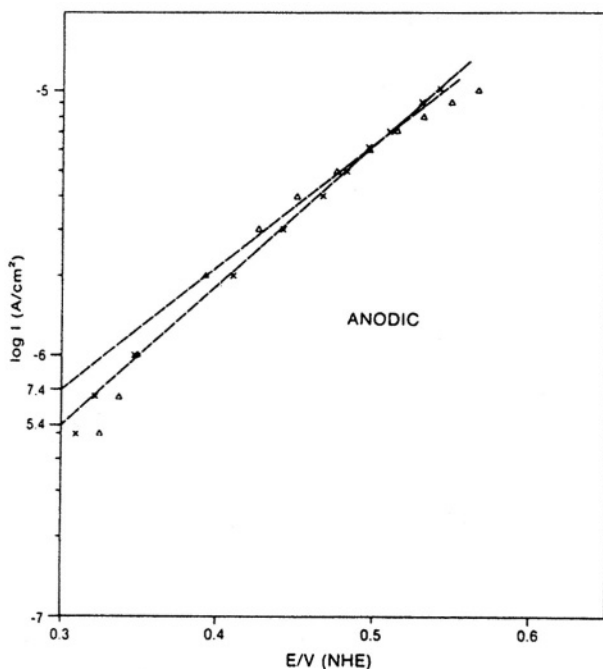
Fig. 14.31. Logarithm of the anodic current density–potential curves of the Q/QH_2 redox couple on gold electrode covered by (×) three layers of DPPC + gramicidin, (△) five layers of DPPC + gramicidin. (Reprinted from A. Rejou-Michel, M. A. Habib, and J. O'M. Bockris, "Electron Transfer at Biological Interfaces," in *Electrical Double Layers in Biology*, M. Blank, ed., Fig. 9, p. 175, Plenum, 1986.)

membranes. Thus, membrane potentials in biology may simply be the analogues of the potentials across the membranes of electrochemical cells.

This thinking leads us to consider the evidence that in some biological processes there is a significant electrical conduction through proteins, sometimes by electrons and sometimes by protons. It seems that this electron–proton conduction is a function of the "doping" of proteins by impurities, arising, for example, from the dissociation products of water. At any rate, because of the small distance over which the currents have to flow through the solid phase, often no more than 50 Å, it is possible to have micro fuel cells in the body that give rise to significant currents, even though the *specific* conductance of proteins when wet is still quite small.

Then, we have been looking in the last few sections at the evidence for actual charge transfer across interfaces in biological systems. The real interface that would
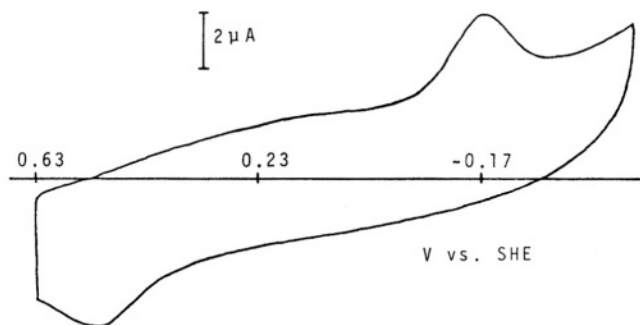
Fig. 14.32. Cyclic voltammogram of coenzyme Q within the bilayer electrode. Phosphate buffer (pH 7.4, ionic strength 0.15), scan rate =100 mV/s. (Reprinted from Y. Xiaoling, J. Cullson, S. Sun and F. M. Hawkridge, "Interfacial Electron Transfer Reactions of heme Proteins," *Charge and Field Effects in Biosystems*, M. J. Allen, S. F. Cleary, and F. M. Hawkridge, eds., vol. 2, p. 87, Plenum, 1989.)

interest us is that between a protein and a solution containing biomolecules. However, among the simplifications of this real biological situation, the most popular experiment involves a study of electron exchange between biomolecules in solution and electrons originating in a metallic or graphite underlay covered with a monolayer of an organic, e.g., bipyridyl. In the absence of this "modifier" of the metallic surface, electron transfer is baulky in such a system or doesn't occur at all. There are two reasons for the poor performance of electron transfer from unadorned metals to biomolecules. In one, the biomolecule, being an enzyme, makes a "crash landing" on the metal and falls to pieces, thereby excluding the possibility of preserving its catalytic properties. Another reason for slow electron transfer to bare metals from the biomolecules adsorbed on it arises from undesired highly resistive layers formed by fragments of the biomolecule.

Apart from these systems, which involve a metal on which is adsorbed a modifier (see above), there is another kind of experiment, although one from which data are as yet less available. One can make up a surface of a metal covered with a biolipid membrane (90% lipid and 10% proteins). There is evidence that some of the proteins in these ensembles are themselves the origin of electrons that can exchange with small redox molecules (e.g., the quinone–hydroquinone system) in solution. Such evidence (though scarce) is significant, for there are no metal underlayers in real biological systems, yet interfacial electron transfer seems to be common there.

The aim in this work is to adsorb enzymes onto electrodes and to have them interact with reduced ions in solution. Some progress is being made in this more advanced area of "enzymes as electrodes." The layer that is adsorbed onto the

electronically conducting underlayer (a single crystal face of gold or of pyrolytic graphite) starts out as a biolipid that can form spontaneously as a self-assembled monolayer (SAM). An enzyme can then be introduced into the biolipid membrane. One now has an enzyme electrode, although the enzyme itself, within the biolipid membrane, may be only a relatively small part, say 10%, of the whole. Some evidence already exists that such an electrode can catalyze electrochemical reactions of quite a complex kind, for example, the oxidation of styrene in solution.

Could one make use of the specificity of enzymes in reacting with specific molecules in solution, which would exist only in a body carrying a certain disease? For example, if a blood sample were made to flow past a test electrode containing, say, 100 pinhead-sized patches, each one of a different enzyme reactive to a molecule characteristic of a specific disease and each connected by individual wiring to an outside circuit, current would flow only from the patch containing the enzyme reacting with its "disease molecule" in the blood. Such a device is a research goal for the twenty-first century.

Further Reading

Seminal

1. W. Nernst, *Z. Physikal Chemie* **2**: 613 (1888). The first application of liquid junction potential theory to explain membrane potentials.
2. J. Bernstein, *Pflüg. Arch.* **92**: 521 (1902). First electrochemical theory involving Na^+, K^+, and Cl^- ions in nerve conduction.
3. J. Bernstein and A. Tchermak, *Pflug. Arch. Ges. Physiol.* **112**: 439 (1906). Differential permeability is the basis to membrane potentials.
4. F. G. Donnan, *Chem. Rev.* **1**: 73 (1924). Selective permeability theory of membrane potentials.
5. E. J. Lund, *J. Zool.* **51**: 265 (1928). Seminal suggestion of electron transfer in biology.
6. Teorell, *Proc. Exp. Biol. Med.* **33**: 282 (1935). Helmholtz layer potential difference contribute to membrane potentials.
7. A. Szent-Gyorgyi, *Nature* **148**: 157 (1941). Seminal suggestion of semiconductivity in biological organisms.
8. K. S. Cole, *Arch. Sci. Physiol.* **3**: 253 (1949). Technique for measuring the spike potential.
9. A. L. Hodges and B. Katz, *J. Physiol.* **108**: 37 (1949). Establishment of importance of Na^+ in outer solution in nerve conduction.
10. A. L. Hodgkin and A. F. Huxley, *J. Physiol.* **116**: 497 (1952). The classical theory of the passage of electricity through the nervous system.
11. T. Teorell, *J. Chem. Phys.* **42**: 831 (1959). Mechanism for the passage of current down nerves.
12. M. Kallman and M. Pope, *J. Chem. Phys.* **32**: 300 (1960). Interfacial electron transfer involving insulators in contact with solutions.
13. W. Mehl, J. M. Hale, and F. Lohmann, *J. Electrochem. Soc.* **113**: 1166 (1960). Electrode processes at interfaces involving insulators in contact with ionic solutions.

14. B. Rosenberg and E. Postow, *Ann. N.Y. Acad. Sci.* **158**: 161 (1960). Electronic conductance in biological organisms distinguished from ionic.

15. E. Goldman, *J. Gen. Physiol.* **27**: 37 (1963). Equation for membrane potentials based on application of the Nernst–Planck equation.

16. S. P. S. Digby, *Proc. Roy. Soc. London* **161**: 504 (1965). Electronic conductivity in crustaceans.

17. F. Gutmann and L. Lyons, *Organic Semiconductors,* Wiley, New York, 1967. The first book to gather and discuss data on electronic conductance in biomolecules.

18. D. DeVault, J. H. Parker, and Britton Chance, *Nature* **215**: 642 (1967). Evidence of tunneling in electronic conductance of bio-organisms.

19. J. V. Howarth, *Phil. Trans. Roy. Soc., London, Ser. B* **270**: 425 (1975). First heat measurements in nerve conduction.

20. T. L. Jahn, *Bioelectrochem. Bioenerg.* **1**: 441 (1976). Tests of the classical theory of membrane potentials.

21. J. O'M. Bockris and M. Schuaib, *Trans. Adv. Electrochem. Sci. Tech.* **13**: 4 (1978). Photostimulated electron transfer to and from photosystem 1 and photosystem 2 from ions in aqueous solutions. (First evidence for a photoelectrochemical mechanism in photosynthesis.)

22. I. Taniguichi, E. Toyosawa, H. Yamaguchi, and E. Yasu Roucki, *J. Chem. Soc.* **102**: 915 (1982). Electron transfer through promoters to dissolved proteins.

23. B. Hille, *Ionic Channels in Excitable Membranes,* Sinauer Associates, Sunderland, MA (1984).

24. A. Rejou-Michel, M. A. Habib, and J. O'M. Bockris, *J. Biol. Phys.* **14**: 31 (1986). Electron transfer from a BLM containing a polypeptide to redox ions in solution.

25. L. J. Boguslavsky, in *Modern Aspects of Electrochemistry,* R. E. White, B. E. Conway, and J. O'M. Bockris, eds., Vol. 18, p. 117, Plenum, New York (1986). Charge transfer at membrane/solution interfaces.

26. M. Blank, *Biochim. Biophys. Acta* **906**: 177 (1987). Excitability in nerve membranes: mechanism.

27. J. O'M. Bockris and F. B. Diniz, *Electrochim. Acta* **34**: 567 (1989). An electrode formulation of the potential difference across an electronically conducting polymer membrane in contact with differing redox species on each side of the membrane.

28. R. Pethig, M. H. Capstick, P. R. C. Gascoyne, and F. E. Becker, Ann. Inst. Conf. I.E.E.E. *Eng. Med. Biol.* **12**: 1 (1990). Protonic and electronic conductance in biological organisms.

29. H. T. Tien, *Electronic Aspects of Membrane Chemistry,* Kluwer, Amsterdam (1991).

30. P. D. Barker and A. D. Mank, *J. Am. Chem. Soc.* **114**: 3619 (1992). Evaluation of dynamics of change in metalloproteins at interfaces.

31. M. Blank, "Electrochemistry of Nerve Conduction," in *Modern Aspects of Electrochemistry,* by R. White, B. E. Conway, and J. O'M. Bockris, eds., Vol. 24, p. 1, Plenum, New York (1993).

32. G. K. Rowe, M. T. Carter, J. Richardson, and R. W. Murray, *Langmuir* **11**: 1797 (1995). Obtaining electrode kinetic parameters from cyclic voltammograms involving proteins.

33. T. M. Nahir, R. A. Clark, and E. F. Bowden, *Anal. Chem.* **66**: 2595 (1996). Linear sweep voltamograms with cytochrome *c* adsorbed on SAMs.
34. Z. Zhang, A. E. Nasar, Z. Lu, J. B. Schenkman, and J. E. Rusling, *J. Chem. Soc. Faraday Trans.* **93**: 1769 (1997). Myoglobin in a BLM and the reduction of chloracetic acid.
35. A. C. Onuoha, X. Zu, and J. F. Rusting, *J. Am. Chem. Soc.* **119**: 3979 (1997). Oxidation of styrene at an interface involving myoglobin in a BLM.

14.6. ELECTROCHEMICAL COMMUNICATION IN BIOLOGICAL ORGANISMS

14.6.1. Introduction

So far in this chapter we have taken some systems from biology (membranes, nerves) and treated them reductively, isolating them in our thinking as much as possible from the complexities of real living systems. Even so, some of the differences between biochemical and chemical systems could not be ignored; for example, proteins are 10–100 times larger than the particles we deal with in normal electrochemistry, although they surprise us by their nimble nature as far as charge exchanges with electrons from and to metals are concerned.

All the time our discussion has been going on, a molecular biologist, listening to it, will have the opinion that what has been gained by the simplifications and isolations we have been making has nevertheless lost something essential. For biological systems should not be considered molecule by molecule or even biological cell by biological cell. The entities all interact. Further, this interaction is much more than the universally known attraction and repulsion between electrostatic charges. The interaction that must be taken into account is often a cell-to-cell interaction so that what one cell "feels" gets communicated to a number of others. Sometimes this interaction extends more than micrometers. For example, in the endocrine system, substances can be released into the blood stream and reach targeted cells far away.

Until the 1980s much of this material on communication in biology was speculative or deduced as likely by interpreting what happened to larger systems in terms of individual cells. However, it is now possible, by the use of ultramicroelectrodes (Section 7.5.4) to get down to the level of detecting chemical changes belonging to an individual cell. Knowing the chemicals that a cell produces and then what happens as a consequence to its neighbor, is enabling us to understand electrochemically something of how cells interact (Fig 14.33).

However, apart from the interaction of a chemical produced in one cell with another, there is an entirely different way in which biological cells communicate; this is by means of electric and magnetic fields. Many of the realizations here are very surprising because it has been found that biological cells are sensitive to electric fields so minute that one has to look very hard at the evidence to sustain belief in its reality.
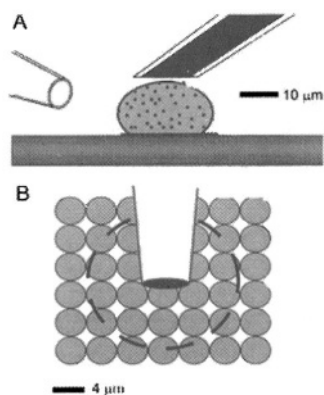
Fig. 14.33. Cartoons of cell–microelectrode experimental arrangements. (A) In a typical single-cell experiment, a glass-encased carbon fiber microelectrode is positioned very close to or touching an individual cell in culture. A glass micropipette positioned nearby is used to administer stimulatory agents to the cell. The electrode can detect electroactive chemical messengers after release from secretory vesicles contained in the cell. (B) For *in vivo* experiments, signals are averaged from a number of release sites. Each dopamine nerve terminal and its surrounding extracellular space occupies approximately 2 μm, illustrated by the small solid circles. Upon release, dopamine freely diffuses from its release site. The dopamine concentration detected at the microelectrode represents the average extracellular concentration in a sphere with an approximate radius of 10 μm, illustrated by the thick dashed line. (Reprinted from R. M. Wightman et al., "Chemical Communication," *Interface,* **5**(3):22, Fig. 1, 1996. Reproduced by permission of the Electrochemical Society, Inc.)

To understand how some of these electrical interactions work, we must look into the extracellular space between cells. We might imagine these spaces as narrow, fluid "gutters" (Adey, 1989), no more than 150 Å wide, in which float hormones, antibodies, and neurotransmitters; and as spaces that are electrically important because of their low impedance compared with the high impedance of the cell membrane.

Some inkling of how cells can be so sensitive to very tiny electric fields can be obtained when it is recalled that thin strands from the helical proteins in the membranes of the cells protrude into the extracellular space. These contain charged receptor sites that are sensitive to electric fields. However, they also themselves generate electric fields and these signals pass to neighboring cells through what is called *protein plaque*, material that forms the junction between cells. If these plaque "connectors" get removed, the intracellular electric signals fail and the cell begins to grow in an unregulated fashion, no longer "in touch" with the rest of the organism. This is one mechanism of carcinogenesis, and because it is basically electrochemical, it is nearer to the type of explanation for cancer Szent-Gyorgyi sought than the purely genetic (or damaged gene) type of cause, which is more frequently cited.

This latter example shows rather well that in living systems, dramatically different things occur if one is concerned with the isolated cell (growing unregulated) or the cell in electrical contact with the rest of the organism (growing according to a pattern).

The bioelectric sensitivities to low-frequency electric fields are shown in Table 14.2. The sensitivity of biological cells to such low fields may have effects far outside electrochemistry. The earth's magnetic field varies with the geographic location. It is significant in navigation for some species of birds. Might it have a role in evolution of species subject to different fields, depending on their location? Can one find here a basis for alleged cancer-causing effects of electric fields from high voltage supply lines or radio towers?

TABLE 14.2
Bioelectric Sensitivities to ELF Fields

	Function	Tissue Gradient	Imposed Field
Sharks and rays	Navigation and predation	10^{-8} V/cm	dc to 10 Hz
Birds	Navigation	10^{-7} V/cm	0.3 G
Birds	Circadian rhythms	10^{-7} V/cm	10 Hz, 2.5 V/m
Monkeys	Subjective time estimations	10^{-7} V/cm	7 Hz, 10 V/m
Man	Circadian rhythms	10^{-7} V/cm	10 Hz, 2.5 V/m
Comparison with intrinsic, cell and tissue neuroelectric gradients			
	Membrane potential	10^5 V/cm	
	Synaptic potential	10^3 V/cm	
	Electroencephalogram	10^{-1} V/cm	

Source: Reprinted from W. R. Adey, "The Extracellular Space and Energetic Hierarchies in Electro-chemical Signaling Between Cells," in *Charge and Field Effects in Biosystems*, M. J. Allen, S. F. Cleary, and F. M. Hawkridge, eds., Vol. 2, Plenum, 1989, p. 266.

14.6.2. Chemical Signaling

Much signaling from cell to cell is chemical. It happens this way: One cell releases a chemical, which is detected by receptors on the surface of another cell, which in turn responds to the chemical. It is the chemical that is the message, passed from cell to cell. The main advance in studying this communication has been the use of ultramicroelectrodes that can be used to measure (sometimes to identify) the signalling chemical. *Exocytosis* is the technical name given to the process by which these chemicals, stored in vessels called *organelles*, are released from the cell and spilled into the extracellular fluid when the organelles fuse with the cell membrane. Once emptied out into the extracellular fluid, the messenger chemicals diffuse within it to the neighboring cells.

What is the nature of these messenger chemicals, which are detected by the use of microelectrodes so small that they can sense the chemicals emitted from a single cell? In view of what has been said earlier about the large size of proteins, the messenger chemicals are relatively small. For example, acetylcholine is a vital part of the neural transport system in mammals. Its structure is

$$
\begin{array}{ccc}
O & & CH_3 \\
\parallel & & \parallel \\
H_3C-C-O-CH_2-CH_2 \mp N-CH_3OH^- \\
& & | \\
& & CH_3
\end{array}
$$
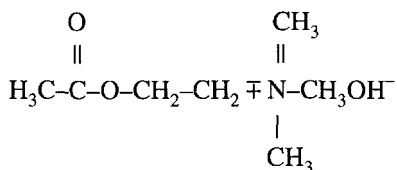
i.e., it is "chemical" rather than biochemical in size.

An electrochemical approach to studying exocytosis is to measure the electrical capacity of the system. A marked increase occurs during exocytosis because of the increase in the area of the system that is active.

The electrode usually used in such experiments is a thin carbon fiber. In order to make repeated fast measurements of successive changes in the substance being secreted, it is necessary to work at an abnormally high sweep rate, around 100 V s^{-1}. Because in a sweep (see Section 8.6) the capacitative current is given by $C(dV/dt)$, then (in spite of the low C for the small area of a microelectrode), this charging current is no longer negligible compared with the kinetic current, which measures exudation of the compound secreted. What is done, then (Miller, 1981), is to take two voltammograms, one in the presence of the organic compound secreted by the cell and one without it. The values of the potentials of the latter curve are subtracted from the former at a series of times. Assuming that the capacity of the system is not affected by contact with the chemical, one can thus obtain thereby a voltammogram free of the interfering $C(dV/dt)$ terms (Fig. 14.34).

There is another way to follow exocytosis experimentally. One potentiostats the ultramicroelectrode and brings it into contact with the cell until the compound it secretes during this process is completely oxidized (Wightman, 1998).
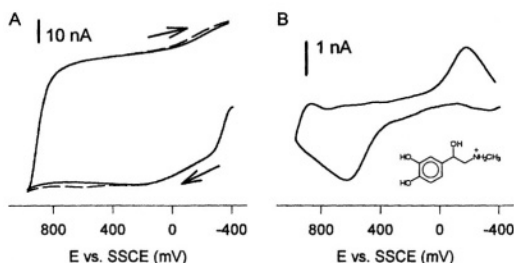
Fig. 14.34. Voltammetry of epinephrine. Background (A, solid line) and signal containing (A, dashed line) currents generated during fast-scan cyclic voltammetry (300 V/s) at a carbon fiber microelectrode $r = 5$ μm). A background subtracted cyclic voltammogram (B) is produced from the traces shown in A. (Reprinted from Wightman, et al. "Chemical Communication," *Interface*, **5**(3): 22, Fig. 2, 1996. Reproduced by permission of the Electrochemical Society, Inc.)

14.6.3. Electrical Signaling

14.6.3.1. *Introduction.* The use of ultramicroelectrodes can do much to detect chemical changes in remote places, including the human brain (Adams, 1976).[12] However, there is also a somewhat different approach to the electrochemical analysis of bio-organisms, and that is to measure their impedances in appropriate frequency ranges. Such an approach was initiated by Cole in 1940 and was vigorously developed by Hermann Schwann at the University of Pennsylvania in the 1960s.[13] The development of impedance analysis in general (Section 7.5.13) has greatly increased the power and scope of the approach in many fields in science (MacDonald, 1983). The approach is particularly good for complex systems because by using a wide

---

[12]Here a statement of appreciation is necessary for Ralph Adams, a professor at the University of Kansas. Already well known for his contributions to analytical electrochemistry, Adams was stimulated to turn his attention in the early 1970s to brain chemistry by his contacts with persons suffering from mental illness. He realized the great importance of being able to "look into" chemical processes within the brain (see Section 14.6.3), and in considering how to do this, made a key discovery: Many of the vital compounds involved were small and easily oxidizable—hence analyzable—at ultramicroelectrodes that could be introduced into the brain without damage. His pioneering work was the nucleus of much that has followed.

[13]Schwann's approach would now be called systems analysis. It was not molecular, but is concentrated upon the electrical properties of integrated systems in specific organisms. Thus, in this kind of approach, the heart might be treated as an oscillating dipole.

frequency range, one can often identify frequency ranges ("windows") that isolate a particular phenomenon in the entire system and pull out information on it.

When a biological organism is bathed in electromagnetic radiation, it manifests at once two regions which, from the impedance point of view, are helpfully different. These two regions are the cell membrane and its extracellular fluid. The resistance of the former is very large and that of the latter very small. This means that assuming the radiation is in the extra-low frequency (ELF) range (e.g., 1–100 cps), the current produced will be almost all in the extracellular fluid.

Enhanced sensitivity in detection of events can be obtained by using the differential plot dZ/dv against $\log \nu$. To demonstrate the kind of results obtained from impedance analysis, Fig. 14.35 shows the impedance variation in the hippocampus of a cat's brain when the stimulus is changed.

14.6.3.2. *Sensitivity of Biological Organisms to Minute Electric Field Strengths.*

In normal electrochemistry, the range of field strengths encountered varies from 1 V cm^{-1} used in measurements of ionic conductance, to 10^7 V cm^{-1} calculated as the field strength in the Helmholtz layer at interfaces. As illustrated in Table 14.2, there is, however, evidence for the response of biological organisms to much, much lower fields, i.e., as low as 10^{-7} V cm^{-1} (Adey, 1981).

How may such minute field strengths have macro effects? This may be made understandable in terms of cooperative action. Consider, for example, the strands of intramembranous proteins that protrude from the phospholipid membrane. They contain charges, and each would react to an applied field. In a membrane at least 10% of the sites will be occupied by proteins. If each protein strand reacts to the field, the very small but coherent signals may be magnified by, for example, 10^{13} strands cm^{-2}. In growing bone cells subjected to a 100-Hz field, say, a significant effect becomes understandable.

14.6.3.3. *Signaling.*

The protein strands referred to in the last section form an inward path from the cell surface to the enzyme systems within the cell. However, in order to find it credible that fields as low as 10^{-7} V cm^{-1} influence the motion of protein strands cooperatively, and that these can electrically transfer a signal of significant magnitude (amplified from the tiny incident signal) to affect active enzymes, one has to overcome the objection that thermal noise is equivalent to a disturbance of kT (or 0.025 eV at 300 °K). Why does this not wipe out the coherence of the alleged cooperative signal? The answer must be speculative. Possibly (Adey, 1989) the cell surface acts like a narrow bandpass filter, accepting the ELF signal but rejecting nearly all the uncoordinated thermal noise.

14.6.3.4. *Carcinogenesis.*

There are many kinds of cancer and many causes. Radiation that is too small in strength to cause ionization in the medium it strikes cannot cause changes in DNA. However, it can interact with material on the surface of cells and hence change—interrupt—the normal signals that these message givers send and that seem to be necessary for regular cell growth. Thus (Trosko, 1989),
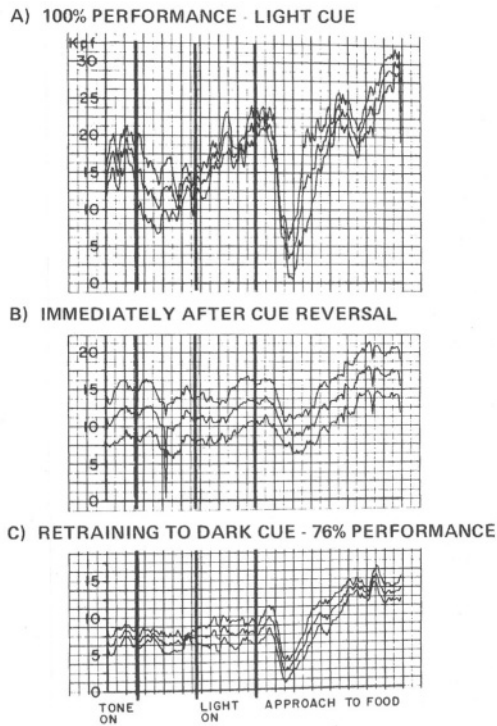
Fig. 14.35. Cat hippocampal impedance averaged over 5-day periods at various levels of training, with successive presentations of alerting, orienting, and discriminative stimuli. In each graph, the middle trace shows the mean. Upper and lower traces show one standard deviation. Variability was low at 100% performance (A), increased substantially after cue reversal (B), but decreased again after retraining (C). (Reprinted from W.R Adey, "The Extracellular Space and Energetic Hierarchies in Electrochemical Signaling Between Cells," in *Charge and Field Effects in Biosystems,* M. J. Allen, S. F. Cleary and F. M. Hawkridge, eds., Vol. 2, p. 265, Fig. 1, Plenum, 1989.)

healthy cell growth requires (1) electrical signals from one cell to another and (2) the transfer of small molecules from one cell to the other. It has been shown experimentally that ELF fields of very low strength can interrupt such signals. Thus, they may be the source of some kinds of cancer. This kind of consideration gives rise to a need for caution in respect to the electromagnetic environment of everyday life.[14]

14.7. ENZYMES AS ELECTRODES

14.7.1. Preliminary

In spite of the early success of Bockris and Schuaib (1978) in registering photocurrents flowing in different directions when the organic compounds of photo-system I and photosystem II were illuminated, in the early 1980s, the prospect of carrying out electron-transfer reactions between proteins immobilized on electrodes and entities in solution seemed unlikely. The most daunting aspect appeared to be the size of the proteins. Figure 14.36 is an attempt to show the relative size of a simple protein, cytochrome *c*. It is instructive to compare the size of this entity with, say, that of O_2, Fe^{2+}, or CH_3OH, the entities involved in oxygen reduction, ferrous oxidation, and methanol conversion to CO_2, respectively. It is by using entities of this latter type that electrode kinetics has been developed; the change to dealing with "molecules" such as that shown in Fig 14.36 is a very big step.

Nevertheless, the possibility of demonstrating the very remarkable specificity of the catalytic power of proteins immobilized on electrodes—to collect or supply electrons from the encounter of enzyme electrodes with entities in solution—looks so attractive that it has spurred investigators (e.g., Bowden, 1993) to efforts that have proved rewarding. The most important advance was made by replacing a bare metal or metal oxide surface with monolayers of organics of a particular type (Fig 14.37) (Eddowes and Hill, 1979).

This interpositioning of specifically chosen organic monolayers (often containing long alkyl chains, sometimes with SH heads) led to what might be called a "soft landing," for proteins upon adsorption; the disaster (Fig 14.20) that occurred when Szucs et al. attempted to adsorb glucose oxidase on Au (decrepitation, dissociation, loss of catalytic power) does not occur if the protein is adsorbed onto a metal (Pt, Au, or pyrolytic graphite) already covered with a layer of "promoter," rather than onto a metal or metal oxide. The fact that electron transfer must occur by tunneling from the underlying metal through the adsorbed promotor to the enzyme—which then reacts with solution components—shows the importance of the proximity of the heme group to the electron source.

---

[14]There is thus a possibility of electromagnetic warfare (Bearden, 1986). A population bathed in ELF radiation at a frequency known to interact with electrochemical processes in the brain might develop a degree of mental retardation.
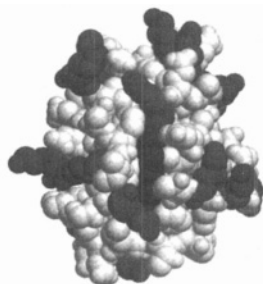
Fig. 14.36. Structure of horse cytochrome *c* obtained from the Brookhaven Protein Data Bank. Shown is the conventional "front face" view highlighting the exposed edge of the heme group (dark gray) located in a region of positive surface charge resulting from several lysine residues. (Reprinted from E. Bowden, "Wiring Mother Nature," *Interface* **6**(4):40–45, Fig. 1, 1997. Reproduced by permission of the Electrochemical Society, Inc.)

Two unexpected results have been generated in the work since 1980[15]:

1. If one were to speculate as to the rate of an electron transfer reaction between a protein and a solution, one might well expect that (because of the giant size and complexity of the species) the rate constant for a redox reaction would be much smaller than that for, say, hydrated Fe^{3+}. The reverse is the case. In fact, in terms of exchange current densities, values as high as 10^{-3} A cm^{-2} s^{-1} can be observed, whereas for the tiny molecules of normal electrochemistry, lower values are more typical.

2. It has been pointed out that cytochrome *c* with its MW of 12,500 is enormous in size compared with the usual entities undergoing electron transfer at interfaces, which have molecular weights of about 100. However, Armstrong et al. (1996) took

---

[15]An interesting issue here is that the use of promoters of different path lengths allows a test of Gamow's equation. (Sec. 9.4.3). Electrons that reach the heme group in cyctrochrome-*c* must tunnel through the adsorbed organic layer, and the length of the alkyl groups that form part of this layer can be varied systematically. The log rate constant is found to be linear with the tunneling distance, *l*, as of course is indicated by Gamow's equation

$$P_T \alpha e^{\frac{-4\pi l}{h} \sqrt{2mE}}$$
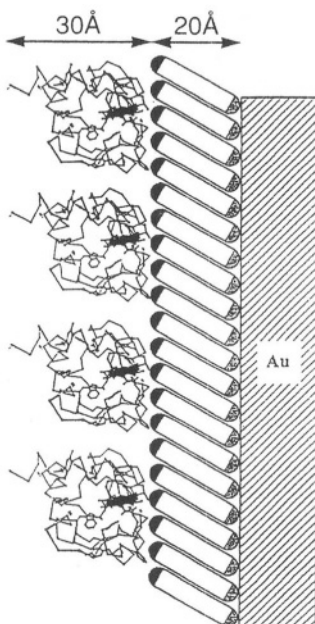
Fig. 14.37. Speculative cartoon of cytochrome adsorbed *c* on a COOH-terminated self-assembled monolayer on gold. A crystalline region for $HS(CH_2)_{15}COOH$ SAM is depicted. Alkanethiol molecules are represented as end-capped cylinders with a COOH terminus (black) and a thiolate (gray) attachment to the gold electrode. Polypeptide line diagrams of cytochrome *c* molecule are shown in an electrostatically favored orientation. For clarity, the heme group has been blackened. (Reprinted from Bowden, "Wiring Mother Nature," *Interface* **6**(4):40–45, Fig. 2, 1997. Reproduced by permission of the Electrochemical Society, Inc.)

much larger proteins, i.e., those for which mere fragments have molecular weights of ~100,000, and carried out so-called "reversible" (i.e., very fast) electron-transfer reactions with them at appropriately structured electrodes. An example is succinate ubiquinone oxido reductase. Here, too, reversible electron transfer can be observed (as long as the complex protein is in contact with a suitable promoter already adsorbed on the electrode to achieve the soft landing). Such complex "real" substances have important *in vivo* activity that offers the possibility of bringing very powerful biocatalysts under *in vitro* control.

14.7.2. What Are Enzymes?

Chemists know that enzymes are catalysts for reactions in the body. In fact, they serve to make the nature of biochemical activities different (several orders of magnitude faster) than the corresponding chemical ones, which are only aided by nonenzymatic catalysts. Enzymes are proteins and their molecular weight varies greatly, from 10^3 to 10^6. Apart from the awesome acceleration of a reaction brought about by enzymes, there is their astounding specificity; they are turned on and off by tiny changes in the structure of the entities taking part in the bioreaction being catalyzed. One must also be familiar with the term "coenzyme," by which is meant another substance, often a vitamin, that has to be present to make the enzyme work. Enzyme catalysts have yet another puzzling feature. Reactions catalyzed by them do indeed increase in rate with temperature. However, proteins tend to unfold at a certain temperature (denaturation) and when this occurs, the enzyme loses its activity. Perhaps that is not surprising, but it is a stimulating experience to learn that each enzyme does this, i.e., it manages to switch off just a little (say, 20 °C above the normal temperature at which the enzyme acts in the organism that produced it.

With all this specificity and the extraordinary power for providing greatly accelerated rates and a close connection to its own bio-organism, the electrochemist must ask if it is possible to take an enzyme out of its natural surroundings, put it on an electrode, have it exchange electrons with the electrode, and still retain its extraordinary activity and specificity for a reaction with a substance in solution that it is meant to catalyze.

14.7.3. Electrodes Carrying Enzymes

The 1980s were spent in discovering what not to do if one wished to carry enzymes on electrodes and use their powers as electrocatalysts:

1. The heme group may not be too deeply buried inside the enzyme. Single electron tunneling jumps greater than about 20 Å do not happen, although much can be done (Heller and Delgani, 1987) to facilitate successive jumps.

2. Even partial decomposition of an enzyme upon adsorption must be prevented because the fragments produced form a passive layer on the electrode and inactivate any enzymatic activity there.
3. Far-reaching decomposition of the enzyme upon adsorption must obviously be avoided, for if it occurs, it destroys the enzyme's catalytic power.

There are only two ways to obtain a successful, adsorbed enzyme acting as an electrocatalyst (Rusling, 1997):

1. The enzyme itself and the electrode system upon which it is to adsorb must be highly purified.
2. The most important point (much applied by Eddowes and Hill at the University of Oxford) is to have a "suitable" promoter on the electrode. It is this prelayer that encourages complicated large biomolecules to adsorb and still retain some of the properties as catalysts that they show in solution.

It has been pointed out (Section 14.2) that one way to make model membranes is to use bilayers of phospholipids placed tail to tail and insert proteins into them (Fig 14.38). In fact, although these arrangements have been described in this chapter as *artificial* entities to be used in laboratories, they also exist in nature, i.e., there are natural membranes with tail-to-tail arrangements of biolipids in which are interspersed the vital protein "wires."

This is the pathway that leads to enzyme electrodes. Thus, Rejou-Michel et al. in 1986 used a Langmuir–Blodgett trough to assemble lipids by gently pushing them
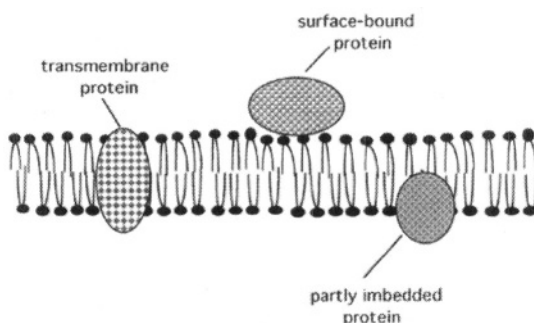
Fig. 14.38. Idealized drawing of a lipid bilayer membrane showing three possible modes of binding for proteins. (Reprinted from H. F. Rusling, "Electrochemical Enzyme Catalysis," *Interface,* **6**(4): 26–31, Fig. 1, 1997. Reproduced with permission from the Electrochemical Society, Inc.)

together, including in them gramicidin (a polypeptide). However, self-assembling monolayers of long-chain compounds can also be formed by taking solutions of long-chain compounds and spreading them on appropriate electrode surfaces, e.g., oriented planes of pyrolytic graphite. Evaporation of the volatile solvent leads to a SAM. An example is given by the work of Rusling and Nasser (1992) in which an electrode covering a BLM film (the analogue of the promoters used by Hill et al.) is made up of didodecyl dimethyl ammonium bromide (DDAB).

didodecyl dimethyl
ammonium bromide (DDAB)   dihexadecyl phosphate (DHP)
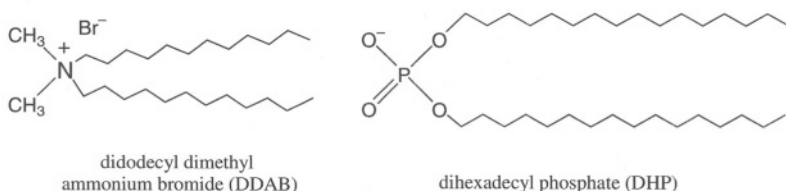
(Reprinted from Rusling, "Electrochemical Enzyme Catalysis," *Interface* **6**:(4) 26–30, 1997, scheme 2. Reproduced with permission of the Electrochemical Society, Inc.)

Into this long-chain electrode-covering substance is put the iron heme muscle protein myoglobulin (MB). (It adsorbs into the SAM from solution.) Now, the point is that when this arrangement is set up, a remarkable increase occurs in the rate of the redox reaction:

$$MBFe\ (III) + H^+ + e \rightleftharpoons MBFe\ (II)$$
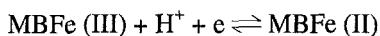
compared with the rate that can be obtained with the enzyme adsorbed on a metal. Figure 14.39 shows a voltammogram in which the cathodic and anodic peaks are separated by only about 0.2 V, an indication of a very fast reaction (see Section 8.6). Here again, the excellent effort of interposing an organic promoter onto the electrode instead of adsorbing the protein directly onto the electrode is demonstrated. (The latter path had been taken by Hawkridge and was indeed successful with In_2O_3 and SnO_2 as the electrode surfaces, but the *rate* of the surface electron-transfer reaction was *several orders of magnitude* slower than that with MB adsorbed into the SAM (Fig. 14.39). Similar reactions have been studied using spinach ferredoxin, a protein active in electron exchanges in chlorophyl and bacteria, and resulting in the production of H.

Why is it that the preadsorbed surfactant layer on the electrode (e.g., the DDAB), has such a helpful effect in facilitating the reactions of enzymes on electrodes? For one thing, the surfactant is a good adsorber on the metal or graphitic electrode. Correspondingly, if, upon adsorption, there is some partial dissociation of the complex enzyme, the preadsorbed surfactant makes it difficult for such fragments to build up passive layers on the electrode, layers that could diminish electron transfer.

Another factor that has to be controlled if one wishes to obtain the maximum reaction rate is pH. At a pH < 4.6, MB is partly unfolded and can be reduced directly. However, at a pH > 9, the reduction involves protonation. The rates of these reactions
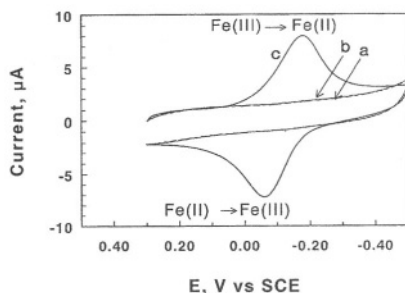
Fig. 14.39. Cyclic voltammograms at 100 mV s^{-1} in pH 5.5 buffer. a, pH 5.5 buffer on bare pyrolytic graphite (PG); b, 0.5 mM MB (horse, from Sigma) in buffer on bare PG; c, MB-DDAB film on PG in buffer, no MB in solution. (Reprinted from J. F. Rusling, "Electrochemical Enzyme Catalysis," *Interface*, **6**(4): 26–31, Fig. 2, 1997. Reproduced with permission from the Electrochemical Society, Inc.)

(remember the huge size of the reactant!) are very high, a rate constant of 8000 M^{-1} cc^{-1}, equivalent to an i_0 of ~10^{-3} A cm^{-2} if the reactant concentration is 10^{-3} M.

14.7.4. The Electrochemical Enzyme-Catalyzed Oxidation of Styrene

Styrene is phenylethylene, $c\text{-C}_6\text{H}_{11}\text{–CH}{=}\text{CH}_2$, a compound important in a number of polymerizations. In the body, the cytochrome P_{450} enzyme in the liver can form styrene oxide, which then may react with the organisms' DNA, i.e., become a carcinogen. The use of MB in lipid films to oxidize styrene has been achieved by Rusling et al. (1997) in a series of reactions that can be written as follows:

$$\text{X-MBFe(III)} + \text{e}^- \rightleftharpoons \text{X-MBFe(II)} \qquad \text{(at electrode)}$$

$$\text{X-MBFe(II)} + \text{O}_2 \rightarrow \text{X-MBFe(II)-O}_2$$

$$\text{X-MBFe(II)} - \text{O}_2 + 2\text{e}^- + 2\,\text{H}^+ \rightarrow \text{X-MBFe(II)} + \text{H}_2\text{O}_2 \qquad \text{(at electrode)}$$

$$\text{H}_2\text{O}_2 + \text{X-MBFe(III)} \rightarrow \text{˙X-MBFe(IV)} = \text{O} + \text{H}_2\text{O}$$

$$2\,\text{˙X-MBFe(IV)} = \text{O} + \text{H}_2\text{O}_2 \rightarrow 2\,\text{HX-MBFe(IV)} = \text{O} + \text{O}_2$$

$$\text{styrene} + \text{˙X-MBFe(IV)} = \text{O} \rightarrow \text{X-MBFe(III)} + \text{styrene oxide}$$
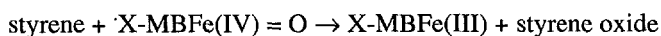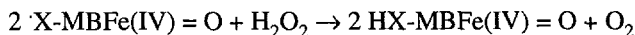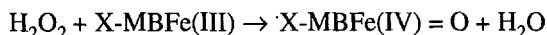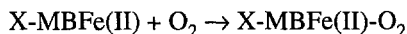
This area is scarcely a decade old. However, it seems reasonable to conclude that the inclusion of enzyme proteins in SAMs provokes fast electrode reactions. It has yet to be determined to what degree it will be possible to retain in the immobilized enzyme the extreme specificity and abnormal catalytic power that it has in its natural state.

14.8. METABOLISM

14.8.1. An Abnormally Efficient Process of Energy Conversion

An internal combustion engine fed with gasoline turns 15–30% of the heat energy released into mechanical power, depending on the manner of working, in particular, the rate. Heat engines, in general, are subject to the Carnot efficiency limit,[16] i.e., the maximum possible efficiency of conversion of heat to mechanical work is $(T_{high} - T_{low})/(T_{high})$, where these temperatures refer to the high and low temperatures that exist in the cycle representing the working of a heat engine.

A mammal, fed with nourishment appropriate to it, converts about 50% of the heat of combustion of food. What is the nature of the energy conversion in biological organisms? There are a number of methods of converting chemical energy to mechanical work, apart from the well-known combustion engine (e.g., thermoelectric), but they all involve a Carnot limit. The T_L for biological conversion in humans would be the normal body temperature, 98 °F or 309 K. To attain an efficiency of energy conversion of 0.5, T_{high} would have to be 618 K or 348 °C, which is overly warm for a biological organism! Nuclear and photodriven processes can be ignored in mammals. The high efficiency of metabolism leads one to the conclusion that its mechanism must involve an electrochemical (fuel cell-like) step, for there the maximum possible efficiency can exceed 0.9 (Bockris and Srinivasan, 1967) and in practice is greater than 0.5.

Although these statements of principle are incontrovertible, the complex nature of biochemical mechanisms demands a more detailed statement than that given before a credible case can be made for an electrochemical mechanism of biological energy conversion. What is generally measured in determinations of metabolism is the food intake in a given time (hence the total heat energy available) and the corresponding mechanical energy expended, for example, on a treadmill. Many things happen to food upon ingestion. It is broken down chemically and converted to a form of energy which

---

[16] Such an expression always yields values in excess of those actually found in the working of an engine. This is because the expression is deduced under the assumption that the engine will work in a thermodynamically reversible way, which is a zeroth approximation kind of statement similar in implication to the statement that an electrochemical cell will perform with zero overpotential at more than 90% efficiency. Of course, in practice, the working of all real heat engines is highly irreversible and the amount of heat energy that can be converted to mechanical work is therefore much less than is indicated by the Carnot limit. Furthermore, the Carnot expression neglects all extrinsic energy losses, e.g., the need to overcome friction and so forth.

then changes adenosine diphosphate (ADP) into adenosine triphosphate (ATP). The latter is a high-energy form of phosphorus and is distributed to where energy is needed in the body, providing the electrostatic energy that is finally manifest in muscular activity with a rejection of ADP. It is the net of all these processes together that is measured when one discusses the efficiency of metabolism. Moreover, the processes that occur on the way from food to muscle are consecutive so that the efficiency being discussed is the product of the various ε_i's. If the number of steps is speculative 12 before that at the mitochrondion/solution interface, and all the ε values are the same, the individual ε of each would be ~0.94. The observed value of ~0.5 is largely the result of the fuel cell-like reactions at the mitochondria and the resultant transfer of energy to ADP.

The title of this book means that our inquiry about metabolism must be limited to the actual energy conversion process itself, and it has been widely agreed that this occurs at the mitochondrion in each of the cells of the organism. It is also agreed that the *distribution* of energy (the currency in respect to wealth) is done by ATP, which yields energy locally when needed.

So, our inquiry into the electrochemical steps in the mechanism of metabolism contains two parts:

1. Is it a tenable idea that RH (representing chemicals from the final products of digestion, such as glucose) is electrochemically converted to electricity at the mitochondrion within biological cells (Fig. 14.41)?

2. Is there an acceptable proposal by which this intermediate electricity (electron flow) developed from two different kinds of sites in the mitochondrion (see Fig. 14.41), is able to convert ADP to ATP?[17] (ATP has to proceed from the mitochondrion to the nearest point at which its energy is needed, e.g., to work muscles.)

14.8.2. Williams Model

An essentially electrochemical model for metabolism and the formation of adenosine tri-phosphate from adenosine diphosphate was suggested by R. J. P. Williams at Oxford University in 1959.[18] It is shown in Fig. 14.40.

Williams has simplified the real activities in mitochondria and made the model look like an H_2-O_2 fuel cell. On the right (i.e., at certain sites in the mitochondrion) there is an anodic dissolution of H_2, which injects electrons into the mitochondrion. At the "other side," at different types of sites in the mitochondrion, O_2 is reduced by

---

[17]One is bound to work at first from analogy. It is clear that the analogue of the present model in the macro world is a fuel cell (the mitochondrion) charging a battery, the essential part of which is the ADP. The mitochondrion is an electrochemical energy converter. It produces electrons at one center and accepts them at another. The ADP to ATP reaction is an uphill reaction, having a positive ΔG°. It has to be *driven*, just as do the reactions in storage batteries during the charging phase.

[18]A competing model called the *chemiosmotic model* was suggested by Mitchell in 1961 and won a Nobel prize. The physical events that which Mitchell's theory implies are less consistent with modern concepts of interfacial charge transfer than those of Williams, which do indicate interfacial charge transfer.
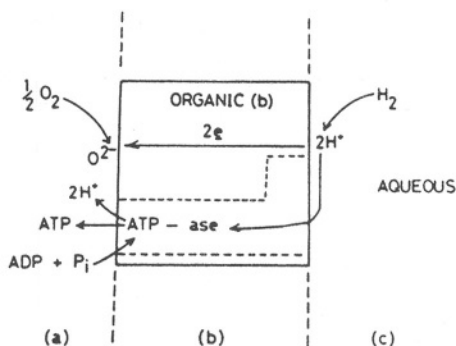
Fig. 14.40. The localized model for ATP formation (in a particle). The particle is illustrated by the box, and the diffusion-controlled paths of e and H^+ are shown. The device can be converted into a chemical model of a membrane [see dashed vertical lines around (b)], to give two aqueous phases (a) and (c) not in equilibrium with (b). This becomes the chemosmosis model by releasing all H^+ to (c) and all O_2^{2-} to (a). These are three models (there are several more) for devices that can use an energized proton to make ATP or to couple other energy-driven processes. (Reprinted from R.J.P. Williams, *FEBS Lett.* **85**: 10, Fig. 1, copyright 1978 with permission from Elsevier Science.)

the electrons coming across from the anodic H_2 dissolution. These events constitute a micro fuel cell, and a potential difference will exist within it as a consequence of the general equation $\Delta G = -nFE$, connecting the free energy change of the net cell reaction to the potential to which it is equivalent.[19] Williams also included in his model (but without elaboration as to how it might take place) the idea that this fuel cell-like conversion of chemical energy to electricity would lead to driving the $ADP \rightarrow ATP$ reaction uphill to store energy in the form of ATP.

---

[19]Writing in 1959, Williams clearly did not take into account the existence of irreversibility and overpotential. The E actually produced in electrochemical energy conversion is less than that given by $\Delta G = -nFE$ because some of the free energy is wasted in overcoming the overpotentials of the two interfacial reactions.

It is not really the dissolved H_2 itself that reacts anodically at the mitochondrion. It is some equivalent hydrogen carrier of biological significance that acts in this way, e.g., RH or reduced nicotinamide adenine dinucleotide (NADH). Correspondingly, the cathodic reaction may not be directly with O_2, but as Szent-Gyorgyi suggested, first with methyl glyoxal (Section 14.3.1). The final reduction of O_2, according to him, is a reaction occurring in the cell plasma.

14.8.3. Development of the Fuel Cell Model in Biological Energy Conversion

Berry et al. (1993) have added significantly to the fuel cell model for biological energy conversion by pointing to the high degree of solid material in biological cells. Thus, a model stressing *surface* reactivity is favored over earlier concepts that were connected with homogeneous chemical reactions in an imagined liquid phase. Another general point arising from discussions of metabolism initiated by Berry et al. concerns the existence in biochemistry of reactions that occur against their free energy gradient. A *chemical* reaction must flow spontaneously in the direction of lowering the free energy of the system. However (cf. the electrolytic dissociation of water), an electrochemical reaction can be *driven* against the spontaneous tendency if there is available an electrical driving force analogous to an outside power source in the working of electrochemical reactors. If some elements of biological systems are in fact equivalent to fuel cells, converting the free energy of chemical reactions to electrical energy, then such elements may indeed drive reactions having a lesser ΔG against their spontaneous tendency.

An electrochemical model of energy conversion in the body, of metabolism, involves the mitochondria within biological cells. These are in contact with a liquid phase and this contains two substances relevant to energy conversion: an H carrier (e.g. NADH), which yields H to take part in the electron-producing ionization $H \rightarrow H^+ + e$, and an O_2 equivalent, which undergoes the electrochemical reduction reaction, $O_2 + 4H^+ + 4e \rightarrow 2H_2O$.

Gutmann et al. (1985) made detailed calculations on a fuel cell model for metabolism. At pH 7 the thermodynamically reversible potential of a glucose-oxygen cell, forming CO_2 and H_2O, would be about 0.5 V. In a human body, the area over which the oxidation and reduction reaction would occur is very large. Estimates of this and of the average rate at which energy conversion must work to consume an input of about 9000 kJ per day for a human organism suggest a mean current density of the oxygen reduction reactions at the mitochondria of about 10^{-6} A cm^{-2}. Enzymatic catalysts of the O_2 reduction may allow an i_0 value as large as this, whereupon the cathodic overpotential can be calculated from Eq. (7.25), i.e.,

$$i = i_0 \frac{\eta F}{RT}$$
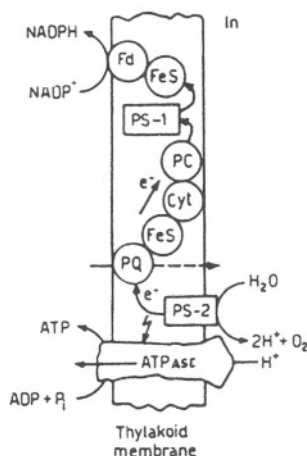
Fig. 14.41. Schematic representation of the biological fuel cell concept. (Reprinted from J. O'M. Bockris and S. U. M. Kahn, *Surface Electrochemistry,* Fig. 7.26, p. 700, Plenum, 1993.)

and with $i_0 = 10^{-6}$ A cm^{-2}, $\eta = 60$ mV at 37 °C. It seems reasonable to take the overpotential for anodic H dissolution as less than that for the O_2 reduction. These considerations, then, led to the model pictured in Fig. 14.40.

Healthy mitochondria would work near the reversible condition. The question of the rate-determining step was considered in the Gutmann et al. model differently from that of Williams, who thought that diffusion of protons through the membranes could be rate determining. In view of the abnormally high diffusion coefficient for H^+ in solution (Section 4.11.3), there seems little likelihood of significant resistance from proton diffusion through membranes, the thickness of which is about only 100 Å. Electrochemical O_2 reduction in real fuel cells is the rate-determining step in energy production and seems to be the likely step here. One of the basic ways in which disease may pull down an organism would be to reduce the i_0 of the oxygen reduction reaction, i.e., to lower the activity of the organism.

14.8.4. Distribution and Storage

Energy distribution and storage in the body are connected with the overall reaction:

$$ADP + H_2PO_4^- + H^+ \rightarrow ATP + H_2O \qquad (14.13)$$
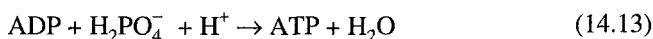
The essence of the electrochemical model (Gutmann and Habib, 1985) is to see this reaction as if it were the overall reaction in an electrochemical cell (occurring at the mitochondrion), presented in Eq. (14.18) in charging mode, going against the free energy gradient. Then, sufficient ATP (the "high energy phosphate") being formed, it is distributed outside the mitochondrial cell to appropriate sites where the "discharge phase of the battery" occurs and energy (the energy of metabolism!) is made available.

How is Eq. (14.18), contrived in electrochemical format, i.e., how does it get driven to the right? In the fuel cell model, this must occur via two constituent reactions just as the overall chemical reaction in a battery is composed of two electrode reactions. Possible reactions are:

$$H_2PO_4^- + H^+ + e \rightarrow HPO_3^- + H_2O \text{ (cathode)}$$

$$ADP + HPO_3^- \rightarrow ATP + e \text{ (anode)}$$
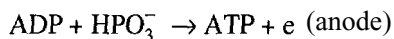
Each would occur at specific but different sites in the mitochondrion. When ATP gives back the energy, these sequences would function in reverse.

The physiology of muscle action and how it is fired are discussed in this book in connection with the mechanism of action of the nervous system (14.4). All that will be said here is that when then above reaction (14.13) run spontaneously in reverse, a supply of energy and protons activates the myofibrils, which are composed of thin filaments made up of the protein actin and thick filaments of the protein myosin. It is the *relative* movement of these two filaments that is the essence of muscle action.

There is evidence from other areas of biochemistry that lends support to the present model. Thus, Rotenberg (1988) correlated the number of cytochrome c sites with the number of ATP synthesis sites. This supports the speculation that cytochrome c can serve as a site on an enzyme electrode, electrodically supplying cathodically active groups for oxygen reduction. Correspondingly, Tsong (1994) has shown that when an anode and cathode are in contact with suspended mitochondrial particles, the rate of ATP formation increases exponentially with the change in potential.

14.9. ELECTROCHEMICAL ASPECTS OF SOME BIOPROCESSES

14.9.1. Introduction

Evidence has been given in previous sections that electrochemical mechanisms play a part in cell stability, biochemical reactivity, the functioning of nerves, etc. It seems therefore reasonable to suspect that some maladies of the body are mediated by electrochemical mechanisms. Examples will be limited to six areas.

14.9.2. Superoxide as a Pretoxin

The electrochemical reduction of oxygen $O_2 + 2H_2O + 4e_0^- \rightarrow 4OH^-$ must be a part of metabolism. The path and rate-determining step depend on the substrate and the pH. However, some mechanisms of the O_2 reduction reaction involve O_2^- (a "superoxide") and it has been suggested (Gerschmann, 1986) that the incomplete consumption of this ion in oxygen reduction allows free O_2^- to accumulate, adsorb on the surface of cells, reach DNA, and cause destruction of part of it, with the resulting cancer, etc.

Sawyer (1988) has reviewed the evidence for this. His conclusion is that if O_2^- is not taken up completely in O_2 reduction, it would be more likely to combine with protons, halogenated carbons, and carbonyl compounds. As a result, peroxy compounds, those containing an O–O bond but having an unpaired electron, would be formed and these would be extremely reactive and in the body, toxic.

It seems worthwhile to explore these ideas further. It would need the investigation of the pathway and rds for the reduction of O_2 on some artificial membranes that simulate suspected reaction centers in the mitochondrion. A pan-electrochemical explanation of degenerative diseases would be to see them in terms of a slowdown of the bioelectrocatalysts of the O_2 reduction. To that extent that age and accumulated fragments reduce the velocity of the enzymatic reduction of O_2, is it likely that an excess O_2 remains left over to form peroxy radicals and damage cells.

14.9.3. Cardiovascular Diseases

Cardiovascular disease is the greatest killer in the United States. There were three theories of its causes current in the late 1990s. The immediate cause is undoubtedly blockage of arteries by a complex mixture of substances that include low-density cholesterol, binding Ca^{2+} and a blood protein, fibrinogen. The theories differ as to what leads to these conditions. The most commonly advocated is that the deposits originate from a diet rich in fats and meat. The second is that they originate from stress (hence the fact that arteriosclerosis is a disease of the technologically advanced countries)[20]; and the third (and most recent) is that homocysteine thiolactone is the critical element; cholesterol carries *it* to block arteries.

It seems strange that cholesterol, a chemical naturally produced in the body and vital to some functions of it, should become the chief killer among technologically active populations in the twentieth century. Arteriosclerosis started to decline in the mid-1960s while the American diet has not undergone a significant change during this time. Further, the Framingham long-term study of the health of a large population over its lifetime has failed to establish a correlation between fat in the diet and high concentration of low-density cholesterol in the blood.

---

[20]Thus, there are reports that a stressful event (e.g., for a surgeon carrying out a difficult operation) can give rise over hours to an increase in total cholesterol.

The major drawback of these classical views is that they fail to take into account the electrical nature of the *stable* constituents in blood (which are colloidal) and that thrombus formation depends upon factors known to control the aggregation of colloids, namely the electric charge, and its sign and magnitude on the colloids in comparison with that on the arterial walls.

Two critical discoveries were made in this direction, largely at the State University of New York's Down State Medical Center. (1) Both the blood corpuscles that form thrombus deposits and the arterial walls in a healthy person are negatively charged, i.e., repelling and there is no thrombus formation. (2) In arteriosclerosis, the sign of one of these charges changes.

Corresponding to these fundamental advances (Srinivasan and Sawyer, 1970), it was discovered that prosthetic materials remain free of thrombi if they are negatively charged and have a potential on the H scale more negative than –0.6 V. This knowledge has significantly influenced the design of prosthetics.

Formation of thrombi (containing cholesterol but also blood proteins and Ca^{2+}) would follow a lessening negative charge on the arterial walls, leading to precipitation. Therefore, the primary cause of the formation of the occlusions that prevent blood flow and cause heart attacks is a reduction in the negative charge on the arterial wall. The corresponding negative charge on the colloidal constituents of blood is unlikely to change.

The arterial wall charge is measured on the assumption that blood plasma is like a dilute electrolyte, so that its zeta potential (Section 6.11.2) reflects this surface charge. It is relatively easy to measure this quantity, and to relate it to the mobility of blood platelets. Figure 14.42 shows the dependence of platelet mobility upon the presence of various chemicals in the blood. Various compounds (above all, heparin but also aspirin) cause an increase in mobility, corresponding to a *more* negative surface charge. Hence, they prevent aggregation and arterial blockage.

These studies have been enhanced by relating the negative charge on the arterial wall to the presence of mucopolysaccharides, which maintain fixed negative charges on the surface while they remain there (Lipinski, 1986). The negativity of the arterial wall is decreased if the blood pH falls below 7.4 and *this occurs if the O_2 content of the blood is insufficient.* Stress releases cationic proteins from blood platelets, which may reduce the negative charge due to the mucopolysaccharides. The same cationic proteins are obtained from red meat.[21]

14.9.4. The Effects of Electromagnetic Radiation on Biological Organisms

Since the early 1980's it has been possible by means of electromagnetic pulses, to bring about the incorporation of drugs into biological cells, to open and shut pores

---

[21]Knowledge of the electrochemical facts of arteriosclerosis shows that the present methods of treating cardiovascular disease reduce the symptoms (the aggregation of colloidal blood chemicals) instead of treating the essential cause, the anionic charge on the arterial walls.
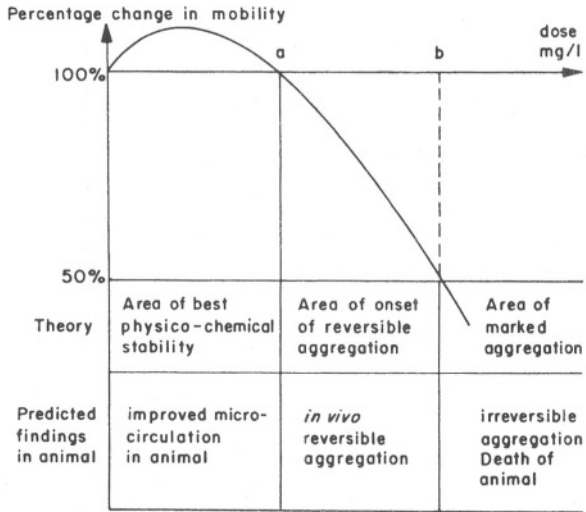
Fig. 14.42. Qualitative representation of percentage change in electrophoretic mobility from control value as a function of concentration of chemical, showing regions of stability or aggregation of cells. (Reprinted from S. Srinivasan, "An Electrochemical Approach for the Solution of Cardiovascular Problems," in *Comprehensive Treatise of Electrochemistry*, Vol. 10, p. 475, S. Srinivisan, Y. A. Chismadzhev, J. O'M. Bockris, and E. Yeager, eds., Plenum, 1985.)

in biological membranes, and to cause certain biochemical reactions. Indeed, these discoveries (Pilla, 1980) constitute an advance parallel to that of Ross Adey's finding that some biological organisms are sensitive to oscillating fields with an intensity as low as 10^{-7} V cm^{-1} (Section 14.6).

Although interesting effects on cell growth have been achieved by subjecting cells to dc electrical fields of $< 10^{-1}$ V cm^{-1} and currents of fractions of microamperes (Becker and Pilla, 1975), as well as ac fields, the most remarkable effects occur when an organism is placed in the magnetic field produced by a Helmholtz coil. The magnitude of the fields used are in the region of 200 μT (microteslas) which is about 10 times the earth's field, with a frequency of about 100 Hz. Such pulsating magnetic fields induce electric currents in the organisms concerned. By this method (Berg, 1989) the growth of yeast has been stimulated (see Fig. 14.43); the production CO_2 from yeast has been triggered; and the stimulation of enzyme activity has been demonstrated (Table 14.3).

Correspondingly (Ramirez, 1998), subjection of human body mechanisms to low-intensity magnetic pulses has been shown to promote physiological changes. Amelioration of osteoarthritic pain for months following an exposure of several hours has been demonstrated.
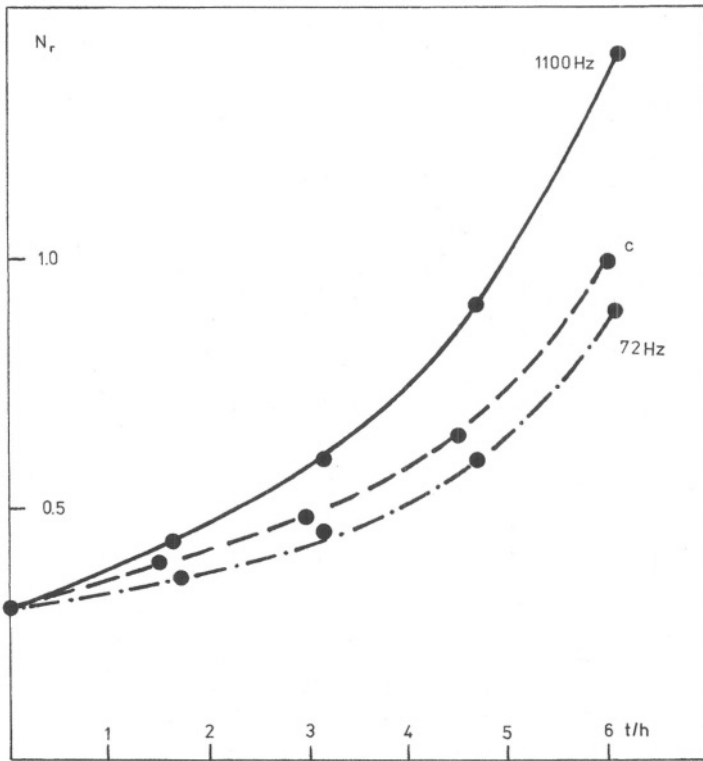
Fig. 14.43. PEMIC influence on yeast proliferation (description in the text). Nr, relative number of colonies; c, control. (Reprinted from *Charge and Field Elects in Biosystems*, M. J. Allen, S. F. Cleary, and F. M. Hawkridge, eds., Vol. 2, p. 310, Fig. 1, Plenum, 1989.)

Such work throws light on accounts of cancer as a result of exposure to the electromagnetic fields generated by electricity supply lines. A field of about 100 V cm^{-1} can be measured near a 60 cps high-voltage power line (Berg, 1989). This magnitude of field produces a corresponding field of 10^{-4} V/cm in human tissue. Correspondingly, television stations working at 100–400 MHz produce fields nearby of about 1 V cm^{-1}. Depending on body orientation with respect to the station, this may produce a 10^{-1} V cm^{-1} field in human tissue!

Now, in laboratory experiments with alternating magnetic fields produced by a Helmoltz coil, the fields induced are up to 10^{-1} V cm^{-1}, and these have been proven to produce biological effects, some of which were described above. These results are important for humans, not only in respect to the effects of power lines, but also their exposure to numerous other sources of electromagnetic radiation, including the wiring

TABLE 14.3
Electrostimulation of Enzyme Activity

Enzyme and Source	Reaction	Effect
ATPase (beef heart)	ATP synthesis	Increase
ATPase (*E. coli*)	ATP synthesis	Increase
Na, K-ATPase (erythrocytes)	K^+, Rb^+ uptake 25 °C	Increase
Adenylcyclase (erythrocytes)	c-AMP	Increase
Adenylcyclase (fibroblasts)	c-AMP (long-term)	Inhibition
Tyrosinase (melanoma cells)	c-AMP, melanin	Increase
Ornithine decarboxylase (fibroblasts)	4–6 hr +phorbolester	Decrease
	10–12 hr + phorbolester	Increase

Source: Reprinted from H. Berg, in M. J. Allen, S. F. Cleary, and F. M. Hawkridge, eds., *Charge and Field Effects in Biosystems*, Vol. 2, Plenum, 1989, p. 301.

of houses, electric blankets, household appliances, etc. In assessing the effects of such results, it is important to realize that the frequency window in which the biological effects occur is quite narrow and also critical. Thus, the detection of a significant field caused by a nearby ac current may not affect biological organisms if its frequency is outside that in which physiological effects have been found (< 100 cps). Conversely, these remarkable effects must be examined also for their positive implications: to what extent will designed electrobiosynthesis be possible?

The effects of fields of low strength were discussed in Section 14.6. There are still controversial aspects of this work, in particular the existence of effects when the strength of the field is less than that of thermal noise.[22]

14.9.5. Microbial Effects

There are two types of electrochemical interactions with bacteria: bactericidal and fuel-cell related.

14.9.5.1. *Bactericidal.* Metallic plates in sea water can become covered with bacterial "sludge," dead bacteria that build up to thicknesses that affect heat conductivity through the metal. Such plates can be protected by controlling the plate potentiostatically in the potential region in sea water in which the reaction

$$O_2 + 2H^+ + 2e \rightarrow H_2O_2$$
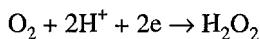
[22]Are the effects discussed in this section electrochemical? They arise because of fluctuating electromagnetic and in some of them, purely magnetic, fields in the body tissue. However, what they induce in the tissues are electric currents, ionic movements, and those and their resulting effects involve electrochemical mechanisms.

takes place. H_2O_2 is an excellent bactericide (Dhar and Lewis, 1981).

Passing a direct current (1 V applied, in milliamperes) through a solution containing bacteria causes the instant demise of between within the current lines (Stoner, 1979). Why do they die under such mild conditions? Is it because of electroporation, in which the solution ions enter the bacteria's cell and separate the contents?

14.9.5.2. Fuel-Cell Related.   It is easy to find bacteria (e.g., *Clostridia butyricum*) that reduce waste materials (including sewage) and produce H_2. An electrode couple consisting of an air electrode separated by a proton-conducting membrane from a compartment containing wastes and *Clostridia butyricum* will produce electricity in the same way as an H_2-O_2 fuel cell. Clearly, some precautions must be taken, such as avoiding the effect of catalytic poisons on the H_2 anode, which reacts with the hydrogen produced by bacteria. However, the economic advantages of producing useful electricity while remediating wastes cleanly could compensate for a certain reduction in efficiency of the fuel cell. The important question, if one considers a commercial process, would be the cost of breeding a sufficient supply of the right kind of bacteria.

14.9.6. Electrochemical Growth of Bones and Related Phenomena

Nineteenth-century records report successful electrochemical healing of broken bones (Stevens, 1812). The beginnings of a modern phase in this work are attributed to Brighton at the University of Pennsylvania (1966). The technique has been developed so that it is an accepted method in orthopedics. The beginning of a noninvasive technique using a Helmholtz coil to induce currents is attributed to Pilla (1974). Both dc and ac currents have been used. Typically, the methods employ pulses lasting 5×10^{-3} s with a repetition rate of 15 per second.

The phenomenon of bone growth under these conditions has as yet no accepted explanation. Ca^{2+} and PO_4^{3-} migrate in a field that has a frequency of 3 kHz. Perhaps the alternating nature of the current causes the formation of apatite [$Ca_5(PO_4)_3$ X where X is OH or a halogen] a constituent of bone, on both ends of the surfaces to be grown together.

Related phenomena have been discovered. For example, osteoporosis can be cured by the application of a direct current, the bone being the cathode. It seems that it is Cu^{2+} and its diffusion out from the bone that causes the osteoporosis, so that "decalcification," which is often the term used to describe osteoporosis, is a misnomer (Becker, 1991).

Dental caries has its electrochemistry, too. The surface charge changes at pH 3.8 and the more acid environment encourages Ca^{2+} to dissolve. Periodontitis ("softening of the gums") arises from bacterial action. However (Van der Kuijii, 1985), if an electrochemical device is introduced into the mouth and the potential of the cathode fixed so that O_2 in the saliva is reduced to H_2O_2, the bacteria and the disease disappear.

14.9.7. Electroanalgesia

The relief of pain by applying electrical wave forms to the affected area has been commercialized in Russia (Persianov and Kastrubin, 1979). Sleep may be induced and it is suggested that the use of electrical wave forms as an anesthetic would arise the after-effects of the more frequently used chemical anesthetics. Typically, two electrodes are placed on the patients' head and two on the neck. Pulses are applied (duration ~0.5 ms) at a frequency of 500–1500 Hz. The resulting tissue current is 1–2 mA. It seems that the use of the method in childbirth is effective not only in eliminating pain but also in dilating the opening of the uterus.

Pulses of 50–125-Hz have been used as an analgesic in dentistry. The mechanistic understanding of these effects (which are related to those of acupuncture) has not been pursued. Perhaps the wave forms are effective in stimulating morphinelike compounds. Alternatively, the applied currents may simply overload currents carrying the natural pain signals to the brain.

14.9.8. Other Effects

The field of electrochemical effects in biology is broad, although up to the end of the twentieth century, the study of their mechanisms had been little pursued. Among the electrochemical effects that have attained the early experimental stage (Findl, 1995) are an increase in enzyme activity, change in shape of DNA with potential when adsorbed on an Hg electrode, acceleration of wound healing, and prevention of bacterial infection.

While the focus of work in the past century has been on the biological and biochemical causes of disease and its treatment, there is another aspect that has received little attention. This is the electrophysiological aspect, and it suffers much from the poor development of theory at a biomolecular level. There is also a further aspect to disease, the psychosomatic. There is much evidence that one can think one's self ill and think one's self well. What is not yet understood is the relation between the electrochemical aspects of disease and the psychosomatic. Are they two sides of the same coin? Should not their study be greatly intensified?

14.10. MONITORING NEUROTRANSMITTERS IN THE INTACT BRAIN AND OTHER SINGLE-CELL STUDIES

14.10.1. Introduction

The use of fast scan cyclic voltammetry has already been described (Section 8.6). In general, microelectrodes, in some cases modified by electrocatalysts, are making it possible to learn about biological events on the scale of a single cell. Among the more important achievements (Wightmann, 1996) is the monitoring of dopamine released after stimulation from neurons in the intact brain and involved in neurotransmission.

The detection of (specifically) dopamine is hindered by the presence in the extracellular fluid of several compounds having redox potentials close to that of dopamine. The technique most likely to succeed here is fast scan cyclic voltammetry (Section 8.6) because the voltamogram provides characteristics that are indicative of the individual compound being monitored. The microelectrodes used have radii of 5 μm, but even this is not small enough to be able to determine dopamine from just one cell. The reacting compounds come from several nerve endings. Nevertheless, the fast scan cyclic voltammetry technique her sufficient time and resolution to allow information to be obtained on the part played by dopamine in neurotransmission in the brain. For example, it answers such questions as: does the released dopamine stay at the synapse or does it diffuse in the extracellular fluid to contact other neurons?

It has been found (Wightmann, 1996) that the lifetime of the dopamine near the originating synapse is short. The dopamine gets used again and is assisted by a protein that helps to transport the dopamine. The total diffusion distance after release is 10 μm. A single release can contact several sites, differing in this respect from the more rigid pathway it follows in other parts of the nervous system.

Although neurotransmitters in the brain offer a tempting area for work, there are other areas in which the detailed knowledge that microelectrode studies can provide
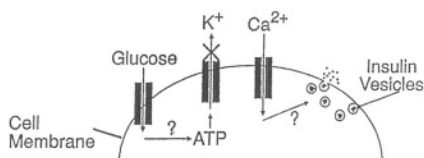
Fig. 14.44. Schematic representation of insulin secretion from single cells. Glucose is transported into the cell by a glucose transporter, where it is metabolized. One of the metabolic products, probably ATP, causes a K^+ channel to close, which leads to cell depolarization. The change in voltage across the cell membrane opens a Ca^{2+} channel, which allows Ca^{2+} to enter the cell and trigger exocytosis of the insulin-containing vesicles by an unknown mechanism (Reprinted from Lan Huang and Robert Kennedy, "Exploring Single-Cell Dynamics Using Chemically-Modified Microelectrodes," in *Trends in Analytical Chemistry*, Vol. 14, p. 159, Fig. 1, copyright 1995, with permission from Elsevier Science.)

may lead to great societal benefit. For example, in spite of the importance of diabetes, many details of the mechanism of insulin release in healthy cells are not yet understood. Figure 14.44 illustrates the factors present in insulin secretion from a single cell.

It would obviously be informative if the release of insulin could be monitored spatially and on a useful time scale. In many microelectrode studies in biology, the substances concerned are easy to reduce or oxidize electrochemically. This is not so with sulfur-containing insulin, and it was not until a suitable (if complex) electrocatalyst was discovered that fast scan cyclic voltammetry with microelectrodes could be used to monitor it. The substance that promotes reaction with the sulfide atoms in insulin is a mixed-valent RuO_2-cyanoruthenate mixture. The catalyst is prepared *in situ* and deposited onto the carbon fiber of the microelectrode. The response time of electrodes thus prepared (Kennedy and Huang, 1995) is < 100 ms, and detection limits are as low as 5 μM.

Such electrodes have been used to examine insulin secretion from single pancreatic β cells. A stimulant is introduced to contact a single β cell adhering to the bottom of a petri dish. The microelectrode is brought into contact with the cell. The result (representing insulin secretion) is shown in Fig 14.45. The peaks shown are Ca^{2+} dependent, and this is characteristic of an exocytotic process (Section 14.10.1). The area under the peak represents 360,000 insulin molecules. The results show that the spikes correspond to the ejection of packets of insulin secreted in exocytosis.
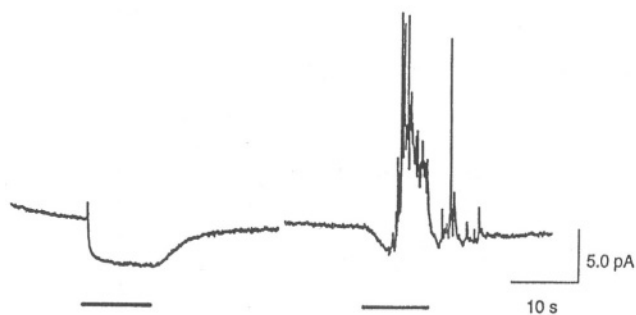
Fig. 14.45. Amperometric recordings from a single canine pancreatic β-cell bathed in Ca^{2+}-free physiological buffer. Bars indicate the application of solution from a micropipette. In the trace on the left, the solution applied was Ca^{2+}-free physiological buffer with 64 mM K^+. (Reprinted from Lan Huang and Robert Kennedy, "Exploring Single-Cell Dynamics Using Chemically-Modified Microelectrodes," in *Trends in Analytical Chemistry*, Vol. 14, p. 160, Fig. 3, copyright 1995, with permission from Elsevier Science.)

Several other detections of substances at the cellular level were published in the mid-1990s, including NO and glucose. All these methods depend upon the good response time and spatial sensitivity made possible by the use of ultramicroelectrodes.

14.11. SUMMARY: MEDICAL EFFECTS, BRAIN, AND SINGLE-CELL EXPERIMENTS

It is easy to summarize the last three sections. The first described an electrochemical fuel cell model for the conversion of the chemical energy of oxidation to electricity and its transfer to ADP.

Several applications of electrochemistry to medical and biological matters were presented. For example, the possibility that O_2^- (a "superoxide"), which is an intermediate in some mechanisms of the reduction of oxygen, could cause degenerative diseases by adsorbing on cell surfaces and decomposing DNA was examined. The idea was turned down, but a related one emphasized; O_2^- can combine with protons and other radicals, and these (e.g., peroxy radicals) are very reactive and may indeed cause trouble.

Another area concerned the number one killer of Americans: cardiovascular disease. In the view most popular at the end of the twentieth century, various blocking materials, centered on low-density cholesterol, form an unwanted lining in arteries near the heart and eventually occlude the flow of blood, causing a heart attack or stroke. From an electrochemical viewpoint, this theory deals with the symptoms but not the cause. Why, the electrochemical view asks, is it that only in some people do the charged components of blood deposit and aggregate on the arterial walls? The answer is based on the experimental determination of the zeta potential of colloidal particles in blood-blocking substances deposited on the arterial wall. The formation of this plaque depends on the electrostatic charge on the surface of the wall. If the charge on the wall is sufficiently negative, no coagulation occurs (for the colloidal material of the blood has a negative charge). The healthy arterial surface contains fixed negative charges of mucosaccharides, and these are unlikely landing fields for the negatively charged particles carried in the blood. However, in sickness and old age, there is a tendency for the strength of the negative charge on the arterial wall to lessen. The collodial particles in the blood coagulate and adhere to the walls, and eventually blood can no longer flow because the artery is blocked with solid material. Much of the deposited, blocking material involves a naturally produced substance, cholesterol, and that is why the therapeutics of arterial disease are aimed at the reduction of cholesterol. However, to scientists who understand the electrochemistry of arteriosclerosis, this approach is similar to treating, say, measles by attempting to eliminate the spots on the face.

The third discussion concerned a field in which at first one has to seek the electrochemistry: the effects of electromagnetic radiation on the body. Here, all is not entirely clear, or, at least, surprising things happen. There is no mystery about the electromagnetic fields that radiate from power lines or anywhere there is a flow of

alternating current, even in the wiring of a house. The surprise is the extraordinary sensitivity of living organisms to electric fields. It turns out that tiny ac fields (10^{-4} V cm^{-1}) have deleterious effects which (when received within some critical frequency window) involve the interruption of cell replication and hint at cancer. But why is it that these effects can be experimentally established when they contain less energy than that of the directionless thermal noise?

Our second section enters deep into electrophysiology. The electrochemistry is here a tool, but a remarkable one in which use is made of electrodes so thin that they can be inserted into the brain without damage. Here, they provide information on one of the neurotransmitters, dopamine, and what happens to it once it has been produced in a pulse from a neuron.

A microelectrode alone is not enough; it has to be used in connection with a fast scan and not only that, but sometimes with an electrocatalyst. This is the case for single-cell monitoring of insulin and a process called exocytosis in which a bunch of insulin molecules, about 360,000 of them, are deposited, eventually to reach the blood stream and to trigger mechanisms that will consume the glucose that has risen beyond a healthy limit. Remarkably, when one considers the cost to the nation of diabetes, all too little is known about it at this molecular level. To craft a cure, here, too, one must look to the microelectrode and its abilities to obtain sufficient knowledge of events at the level of the individual cell.

Further Reading

1. R. J. P. Williams, *The Enzymes,* Vol. 1, p 391, Academic Press, New York (1959). A first statement of Williams' theory of metabolism.
2. J. O'M. Bockris and S. Srinivasan, *Nature* **215**: 397 (1967). Only possible to explain high efficiency of metabolism if energy conversion has fuel cell mechanism.
3. P. N. Sawyer and S. Srinivasan, *J. Coll. Interface Sci.* **32**: 456 (1970). Coagulation of biomaterials in blood as a function of the surface charge.
4. S. Srinivasan and P. N. Sawyer, *J. Coll. Interface Sci.* **32**: 456 (1970). The stability of prosthetic material as a function of its surface charge.
5. R. O. Becker, *Nature* **235**: 109 (1972). The stimulation of cell growth under weak electric fields.
6. R. N. Adams, *Anal. Chem.* **48**: 1126A (1976). Investigation of the electrochemistry of the brain.
7. A. Pilla, C. A. Basset, S. Mitchell, and L. Norton, *Acta Orthopaedic Belg.* **46**: 700 (1978). Stimulation of bone growth.
8. J. S. Clegg, in *Water Structure in Cell Associated Water,* W. Drost-Hansen and J. Clegg, eds., p. 363, Academic Press, New York (1979).
9. J. O'M. Bockris and M. Tunulli, *J. Electroanal. Chem.* **100**: 7 (1979). Bioelectrochemical energy storage mechanism.
10. M. V. Berry, *FEBS Lett.* **117:** (Supplement) K106 (1980). Enzymes in cells are adsorbed on cell surfaces.

11. J. O'M. Bockris, F. Gutmann, and M. A. Habib, *J. Biol. Phys.* **13**: 31 (1985). A fuel cell mechanism in biological energy conversion.

12. R. Gerschmann, D. Gilbert, S. Nye, P. Dwyer, and W. Fenn, *Science* **119**: 623 (1986). O_2^- as a general cause of disease.

13. D. Sawyer, *Chenteck.* **18**: 369 (1988). O_2^- is a pretoxin.

14. H. Berg, in *Electromagnetic Fields and Biomembranes,* M. Maikov and M. Blank, eds., Plenum; New York (1988).

15. K. Pikel, T. J. Schooeder, and R. M. Wightman, *Anal. Chem.* **60**: 1268 (1988). Use of microelectrodes to investigate processes in the brain.

16. D. Van der Kuiji, P. A. Vingerling, P. S. Smitt, K. de Groot and J. de Graaf, *Electric Stimulation of Bone Growth,* Karger, New York (1993).

17. P.A. Garris and R. M. Wightman, *J. Neurosci.* **14**: 462 (1994). Fast scan voltammetry and brain electrochemistry.

18. L. Huang and R. Kennedy, *Trends Anal. Chem.* **14**: 158 (1995). Exploring single-cell dynamics: insulin at single cell level.

19. R. M. Wightman, S. Hocksteter, B. Michael, and E. Travis, *Interface* **5**: 23 (1996). Following dopamine in the brain.

EXERCISES

1. In an experiment involving artificial changes in the K^+ and Na^+ inside and outside the axon of the giant squid, the K^+ concentration in the extracellular fluid is increased ten times. Using data in the chapter on the normal concentration of Na^+ and K^+ in the resting state and the resting potential usually observed, calculate by means of the Goldman equation the expected change in the transmembrane potential. (As a simplification, assume that the coefficients multiplying the concentration terms can be equated to the respective ionic diffusion coefficients, i.e., $1.33 \ 10^{-5} \ cm^2 \ s^{-1}$ for Na^+ and $1.96 \ 10^{-5}$ for K^+.) (Bockris)

2. The external and internal ion concentrations for the *Aplysia* giant nerve are as follows:

	extra	intra
$[K^+]$/mM	10	280
$[Na^+]$/mM	485	61
$[Cl^-]$/mM	485	51

The ionic permeabilities for the resting membrane are as follows: $P_K:P_{Na}:P_{Cl} = 1:0.12:1.44$. Calculate the resting membrane potential. (Contractor)

3. The ionic concentrations for the Aplysia giant nerve are the same as in exercise 2. The resting membrane potential is –49 mV. Predict the direction of spontaneous movement of ions in the resting condition. (Contractor)

4. Dry proteins are nonconductors but exhibit semiconductivity when wet. The magnitude of this varies greatly. Some wet proteins exhibit specific conductances on the order of 10^{-4} S cm^{-1}. If energy consumption in the body has an electrochemical mechanism, the corresponding model suggests that the average current density at a mitochondrial membrane is on the order of microamperes cm^{-2} (corresponding to an energy consumption of 10^4 kJ per day). On the basis of this approximate estimate, show that the ohmic potential loss across a biological membrane having the specific conductance mentioned and a thickness of 5 nm would be less than 1 mV. (Bockris).

5. Szent-Gyorgyi theorized that an important step in metabolism was the mediated reduction of O_2. Thus, in this view, O_2 does not undergo interfacial electron transfer, this is the function of methyl glyoxal. The reduced form of this aldehyde is then oxidized homogeneously by O_2, which thus undergoes the required reduction. (a) Write out a reasonable scheme for this idea. Szent-Gyorgyi thought that a diminution in the concentration of methyl glyoxal would interrupt the metabolic pathway and could be a cause of cell death and a pre-cancerous state. (b) If a speculative exchange-current density for the alleged methyl glyoxal reduction step is taken as 10^{-5} A cm^{-2}, calculate the reduction in the concentration of methyl glyoxal necessary to reduce the metabolic rate by 50%. (Bockris)

6. A (more modern) approach to the membrane potentials observed in biology is to take account not only of the liquid junction (Nernst–Planckian) potential aspects, but also to model the net potential difference across the membrane as a bielectrode. On each side of the membrane it is supposed that (differing) electron-transfer reactions occur. The observed potential is the difference of these, plus IR components, of active electronic and ionic potential differences across the membrane.

    Using an expression (low field approximation) for a membrane potential derived from this model calculate the membrane potential assuming that the difference of the reversible redox potentials of the reactions on either side of the membrane is 120 mV; $(i_0)_1 = 10^{-4}$ A cm^{-2}; $(i_0)_2 = 10^{-6}$ A cm^{-2}; real area = 10 times the apparent area (0.2 cm^2); and $R = 10^4$ ohms per sample of membrane. (Bockris)

7. Gutmann et al. suggested that the essential power-producing processes of metabolism consist of the two electrochemical reactions, $O_2 + 4H^+ 4e \rightarrow 2H_2O$ and NaDFH $\rightarrow$ NaDP$^+$ + H$^+$ + 2e. (a) Consider whether this model can be made consistent with the facts of the energetics of metabolism (3000 kcal/day converted to mechanical energy at a 50% conversion efficiency), making plausible assumptions about the surface area in the energy-producing mitochondrial cells of humans. (b) Taking the potential of the proposed fuel cell as 0.38 V, calculate the levelized total watts being used to "run" a human being and estimate the

order of magnitude of the current density at a typical energy producing cell. (Bockris)

8. (a) Explain the meaning of BLM and state how a BLM may be prepared and what it is used for. (b) What is an ionophore and what is it used for? (c) How is "active transport" rationalized in the classical theory of the passage of electricity through nerves? Some toxic effects are known to work through action on the nervous system. (d) By what model could this be rationalized? (Bockris)

PROBLEMS

1. Hill and, independently, Hawkridge, made considerable progress in examining electrode reactions involving protein and enzymes. The key to the success they achieved was the covering of the electrode with what are called "modifiers," for example: 4-4′-bipyridyl. Figure P14.1 contains a number of cyclic voltammograms involving the reactions of redox proteins in the edge plane of graphite. (a) Using knowledge gained from the material in Chapters 8 and 9, discuss these voltammograms with the aid of the information in the caption.

    (b) The order of magnitude of the peak current densities in the diagrams given is $10 \ \mu A \ cm^{-2}$. (b) Assume the heme group is in the center of these proteins and using information on protein structure obtained outside this chapter, estimate the path length of the electron passage from the electrode (on which is 4-4′-bipyridyl is adsorbed) to the heme in the adsorbed protein. (c) Discuss this in terms of the probability of the electron passage over this distance using the Gamow equation (Chapter 9) and the assumption that the energy term is 3 eV. (d) Determine by calculation if the current density mentioned is consistent with a quantum mechanical step from the electrode to the heme. (Bockris)

2. It has been suggested by D. Sawyer that the reduction reaction $O_2 + 4H^- + 4e \rightarrow 2H_2O$, which occurs in metabolism would have, under some circumstances, an incompletely converted intermediate, O_2^-. This so-called superoxide ion could react with protons and other moieties to form peroxy radicals, and these are toxic. Peroxy radicals are thought to adsorb on the DNA of cells and cause their destruction, i.e., act as a trigger to a precancerous state. Consider this in respect to the dose-response relation in cancer caused by radiation. Suppose the same (low) dose (i.e., total number of particles from radioactive substances that cause biological damage) is delivered slowly (low intensity) or at a high intensity. Would the model indicate a greater damage for the low- or high-intensity delivery of the same dose? Fully justify your answer. (Bockris)

3. The best-known theory of the spike potential produced in an impulse down a nerve is that due to Hodgkin, Huxley, and Katz (1952). Although this theory is still current among electrophysiologists, it is now regarded with skepticism by a number of physical scientists who have examined it in the light of modern

Spinach ferredoxin
[2Fe-2S]$^{2+/1+}$

Clostridium ferredoxin
2[4Fe-4S]$^{2+/1+}$

Clostridium rubredoxin
Fe$^{3+/2+}$

Spinach plastocyanin
Cu$^{2+/1+}$
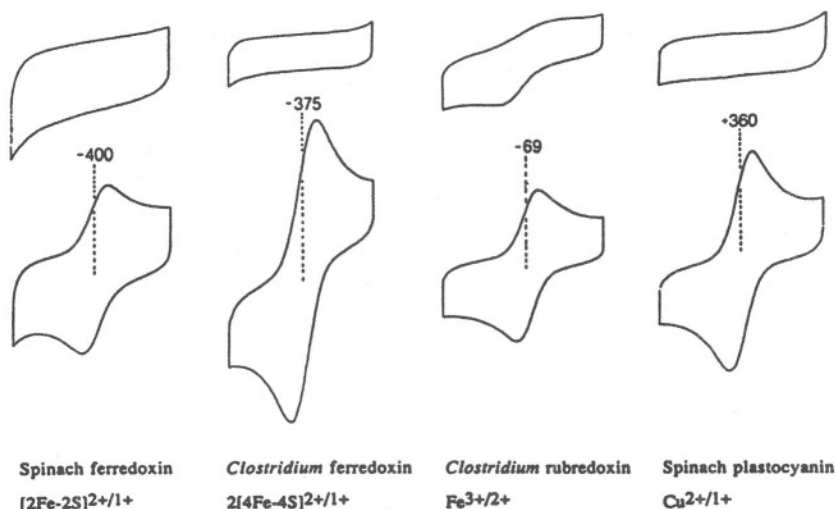
Fig. P14.1 Cyclic voltammetry of a range of small, negatively charged redox proteins at pyrolytic "edge" electrodes, showing the promotion of direct electrochemistry by multivalent cations. Scan rate is 20 mV s^{-1}. Electrodes polished with 0.3-μm alumina slurry and sonicated. Half-wave potentials in millivolts vs. NHE indicated in each case. Spinach [2Fe-2S]ferredoxin: (Top) Protein 65 μM in 5 mM Tricine, 1.0 M NaCl, pH 8.0, temperature 25 °C. Note the absence of response in the 1 :1 electrolyte even at this high concentration. Range is −145 to −680 mV (vs. NHE). (Bottom) Cyclic voltammetry of the same solution after inclusion of 1.5 nM [Cr(NH$_3$)$_6$]$^{3+}$. The latter is *not* redox active in this potential range. *Clostridium pasteurianum* 2[4Fe–4S]ferredoxin: (Top) Protein 150 μM in 5 mM Tricine, 1 mM NaCl, pH 8.0, temperature 25 °C. Range is −106 to −606 mV (vs. NHE). (Bottom) Cyclic voltammetry of the same solution after inclusion of 15 mM decamethonium (DM^{2+}) bromide; DM^{2+} = [(CH$_3$)$_3$N(CH$_2$)$_{10}$N(CH$_3$)$_3$]$^{2+}$. *Clostridium pasteurianum* rubnedoxin: (*Top*) Protein 67 μM in 5 mM Tricine, 0.10 M NaCl, pH 8.0, temperature 25 °C. Range is +194 to −361 mV (vs. NHE). Note that this protein gives a poor but discernible response in this medium (a similar observation is made for *Clostridium pasteurianum* 2[4Fe–4S]ferredoxin in 0.1 M NaCl). (Bottom) Cyclic voltammetry of the same solution after inclusion of 4.4 mM [CA(NH$_3$)$_6$]$^{3+}$. Spinach plastocyanin: (Top) Protein 200 μM in 10 mM MES, 1 mM KCl, pH 6.0, temperature 25 °C. Range is +619 to −175 mV (vs. NHE). (Bottom) Cyclic voltammetry of the same solution after inclusion of 0.5 mM [Pt(NH$_3$)$_6$]$^{4+}$. Reprinted with permission from Armstrong et al. Accounts of Chemical Research, **21**:407, Fig. 1, copyright American Chemical Society.)

knowledge of electrode processes involving semiconduction and insulators.

The magnitude of the spike potential is around a hundred millivolts. (a) Following the spirit of the H-H theory, use the Goldman equation to calculate the change in ionic concentration between the outside and inside of the axon that is needed to explain this spike. (b) From this result, calculate the flux of cations per square centimeter of membrane that would have to pass across the membrane to bring about the concentration change.

In the Goldman equation use the diffusion coefficients in question 1 in place of permeation coefficients. The rate of permeation is $-D(dc/dx)$. The dx is essentially the membrane thickness and its neglect will cancel out of the equation (cf. the Goldman equation). (c) Is this flux consistent with actual radiotracer measurements of the movement concerned? (In the Hodgkin–Huxley and Katz work, arbitrary values were used for the P's to ensure that the equation replicated the experiment. This, of course, makes it difficult to check its validity. Assume the starting concentration of ions on either side of the membrane is that shown in the text. The average internal diameter of a squid axon is about 1 mm.)

Transport of ions across a membrane in the direction of the gradient of the electrochemical potential is called *passive.* However, transport occurs *against* the gradient of the electrochemical potential (active transport). (d) How can this be? (Bockris)

4. Show that the membrane potential, V_m, is given on a Nernst–Planck basis by:

$$V_m = \frac{RT}{F} \ln\left[\frac{P_k[K]_e + P_{Na}[Na]_e + P_{Cl}[Cl]_i}{P_k[K]_i + P_{Na}[Na]_i + P_{Cl}[Cl]_e}\right]$$

(Contractor)

5. Pictures of the blocking of arteries by the buildup of plaque are familiar sight in the offices of cardiologists. The mechanism of this mode of premature death among Americans is a subject of controversy, the generally cited cause (excess low-density cholesterol) being regarded increasingly as a symptom rather than a cause. The latter is to be sought in factors that reduce the negative charge on the arterial walls (Section ??).

Answer the following four questions connected with this important bioelectrochemical topic:

(a) What is the sign of the surface charge of platelets in arterial blood? (b) What would you expect in respect to the variation of the charge on the arterial surface with age? The cause of arteriosclerotic plaque is usually described in terms of cholesterol buildup. (c) How can this explanation be made consistent with the electrochemical one suggested in this book? (d) Should surgeon's instruments be connected to a power source that would charge them negatively?

6. A gold, quartz crystal electrode has been modified to allow the electrode to support a lipid bilayer membrane that houses the enzyme cytochrome c oxidase. This enzyme undergoes a reaction with cytochrome c in solution which can be monitored amperometrically (current plotted as a function of time). This reaction is carried out in a flow-injection cell using a wall-jet configuration. The equation for a theoretical limiting current under wall-jet conditions is as follows: $i_t = 0.898nFCD^{2/3}v^{-5/12}a^{-1/2}A^{3/8}U^{3/4}$ where n = number of electrons (1 in this case), F = Faraday's constant (96,500 C/mol), C = concentration (mol/cm^3), D = diffusion coefficient (1.12×10^{-6} cm^2/mol for cytochrome c), v = kinematic viscosity (0.01 cm^2/s), a = diameter of input conduit (cm), A = electrode area (cm^2), and U = average volume flow rate (cm^3/s).

   (a) If the average limiting current (i_t) obtained for the above reaction is 20 nA, what percentage of the electrode is active, i.e., contains cytochrome c oxidase? The total electrode area is 0.2 cm^2, the concentration of cytochrome c is 10 μM, the flow rate is 0.5 mL/min, and the diameter of the input conduit is 0.072 cm. (Hint: Make sure that the units are correct.) (b). What is the theoretical limiting current (i_t) predicted for the reaction concentration of cytochrome c is 37.5 μM assuming the active area of the electrode calculated in (a)? (Rhoten)

7. Cytochrome c oxidase can be immobilized in an electrode-supported lipid bilayer membrane. The area of the electrode used is 0.2 cm^2, and the diameter of the oxidase is 80 Å. This enzyme is electroactive and yields the cyclic voltammetric response shown in Fig. P14.2.

   Integrating the area under this peak yields the charge passed in the anodic scan. This value can be directly correlated to the amount of oxidase immobilized on the surface. (Assume that the charge passed in this scan results only from a Faradaic process.) If 50% of the electrode contains immobilized oxidase, calculate the charge passed in the above anodic scan. (Hint: Complete oxidation of 1 mol, of oxidase yields 4 mol of electrons.) (Rhoten)

8. The temperature dependence of the kinetics of myoglobin has been studied using cyclic voltammetry. These experiments were conducted over a range of temperatures (10–50 °C) to determine the change in reaction center entropy
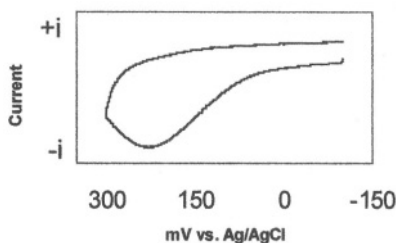
Fig. P14.2

upon electron transfer. Table P.1 contains temperatures and the corresponding $E^{\circ\prime}$ values for the Fe^{3+}/Fe^{2+} redox couple of myoglobin. Calculate the reaction center entropy change from these data.

TABLE P.1

Temperature (°C)	$E^{\circ\prime}$ (volts)
10	−0.175
20	−0.179
30	−0.186
40	−0.188
50	−0.190

9. Cyclic voltammetry is used to study the redox properties of a particular electrochemical system. The usual method has been used to determine the heterogeneous rate constant at various scan rates. These data are shown in Table P.2.

TABLE P.2

Scan Rate (mV/s)	Heterogeneous Rate Constant
10	0.083
20	0.080
50	0.075
100	0.068
200	0.054
500	0.047
10	0.012
20	0.014
50	0.017
100	0.022
200	0.030
500	0.037
10	0.033
20	0.038
50	0.034
100	0.032
200	0.037
500	0.036

(a) Which set of experiments contain reliable data? (b) What do the trends in these data reveal about the experiment? (c) Should the heterogeneous rate constant be dependent or independent of scan rate? (Rhoten)

10. A spectroelectrochemical experiment can be conducted at an optically transparent electrode (OTE) to determine the diffusion coefficient of a species. Single potential step chromoabsorptometry experiments are conducted so that the diffusion

coefficient of cytochrome c may be found. This experiment consist of a potential step perturbation monitored by simple optical absorption spectrophotometry as a function of time. Cytochrome c(II) and cytochrome c(III) have differing absorption maxima. Therefore if cytochrome c(III) is initially present in solution and a reducing potential is applied, the formation of cytochrome c(II) can be monitored spectrophotometrically at a wavelength of 550 nm. The following equation can be used to calculate the diffusion coefficient of the species of interest:

$$\Delta A = \frac{2\Delta\varepsilon_o^{1/2}C_o t^{1/2}}{\pi^{1/2}}$$

From the data given in Table P.3, calculate the diffusion coefficient for cytochrome c. [Assume that the concentration of cytochrome c(III) is 97.6 μM and the difference in molar absorptivity ($\Delta\varepsilon$) is $21,100\ M^{-1}\ cm^{-1}$.] (Rhoten)

TABLE P.3

Time (seconds)	Change in Absorbance
0	0
5	0.0055
10	0.0078
15	0.0095
20	0.011
25	0.012
30	0.013

MICRO RESEARCH PROBLEM

1. It is desirable for diabetics to carry wristwatch meters that make the concentration of glucose in the blood available at a glance. Among the problems to be solved is immobilization of the enzyme glucose oxidase on a microelectrode made, e.g., of pyrolytic graphite. (a) How could this be achieved? (b) What would be the minimum tip diameter of such an electrode? (c) Would its insertion through the skin be feasible? In order to power a readout of the glucose concentration, the enzyme-catalyzed oxidation would have to result in electron transfer through the enzyme to an underlying conductor leading to an amplifier and readout. (d) Perform design calculations to examine (for glucose oxidase) whether such an electron transfer through the enzyme would be a feasible path toward a wristwatch meter for diabetics. (e) By what means could the electronic conductance of the enzyme be artificially increased? (Bockris)

CHAPTER 15

ENVIRONMENTALLY ORIENTED ELECTROCHEMISTRY

15.1. THE ENVIRONMENTAL SITUATION

Although a fossil fuel-based industrial economy had been in full swing for more than a century, it was not until the early 1970s that the air pollution arising from it became a political issue. A wakeup call on environmental pollution had been given in the book by Rachel Carson, *Silent Spring* (1962). It brought out evidence that the absence of birdsong in the spring arose from the birds having eaten dead insects containing DDT.

The 1970s, too, was a time in which an older prophecy (Tyndall, 1861; Plass, 1956) was fulfilled, i.e., that CO_2 from fossil fuel combustion would add significantly to the CO_2 naturally in the atmosphere. The temperature of the earth is largely controlled by the balance between absorption and reflection of solar radiation. The light contains infrared (IR) wavelengths and this kind of light is absorbed by CO_2 and degrades to heat. A greater CO_2 content in the atmosphere (Fig. 15.1) means a higher temperature (the greenhouse effect).

The reduction of visibility during mist and fog is a natural event. However, in the nineteenth century, sulfurous fogs became common in industrial cities, and in the middle twentieth century a new phenomenon appeared, particularly in Southern California: photochemical smog. It turned out that a vital intermediate in the formation of smog, peracetyl nitrate (the smog molecule), came from automotive NO_2 and unsaturated hydrocarbons which led, after a complex series of reactions with photochemically produced NO, to an enervating and visibility-impairing suspension, smog.

Acid rain is another plague that became apparent to the public in the 1980s. It has emptied lakes of their fish, destroyed forests, and, with more immediate economic

Carbon dioxide levels in the atmosphere are increasing exponentially

Concentration, ppm (volume)

1958 60 62 64 66 68 70 72 74 76 78 80 82 84 86 88 90 92
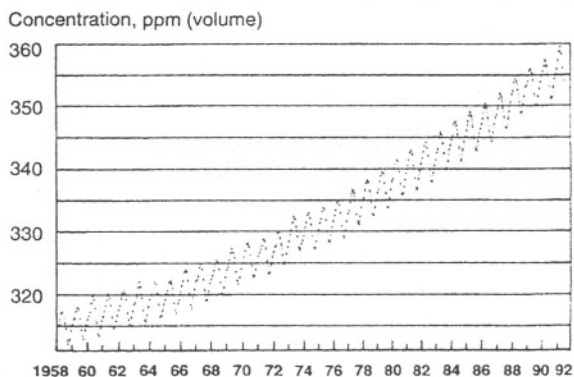
Sources: C. D. Keeling, R. B. Bacastow, A. F. Carter, S. C. Piper, T. P. Whorf, M. Heimann, W. G. Mook, D. J. Moss, and H. Roel-offzen, Scripps Institution of Oceanography, National Oceanic & Atmospheric Administration

Fig. 15.1. Concentrations of carbon dioxide were measured with a continuously recording nondispersive infrared gas analyzer at Mauna Loa Observatory, Hawaii. *The dots indicate average monthly concentrations.* (Reprinted with permission from B. Hileman, "Carbon dioxide levels in the atmosphere are increasing exponentially," *Chem. & Eng. News,* April 27, 1992. pp. 7–19. Copyright 1992 American Chemical Society.)

consequences, attacked the stability of certain building materials. The SO_2 that turns into H_2SO_4 in rain is a by-product of burning coal for electricity (coal contains 3–6% sulfur).

These three environmental problems are the most pressing today, but they are accompanied by a host of others, some of them realized only in the 1990s. One is disposal of household and industrial wastes.

Another health hazard is associated with the ingestion of toxic compounds in foods or their gradual contamination of groundwater. Methyl mercury in fish is the standard example. It has been proven that mothers eating fish containing polynuclear hydrocarbons produce children of reduced intelligence. This exemplifies a very general danger, the dimensions of which we cannot yet know (a high proportion of the population has traces of DDT in the brain). Which toxic organics are being spread far and wide over the population? What long-term health effects will they have?

State of the world

Material standard of living
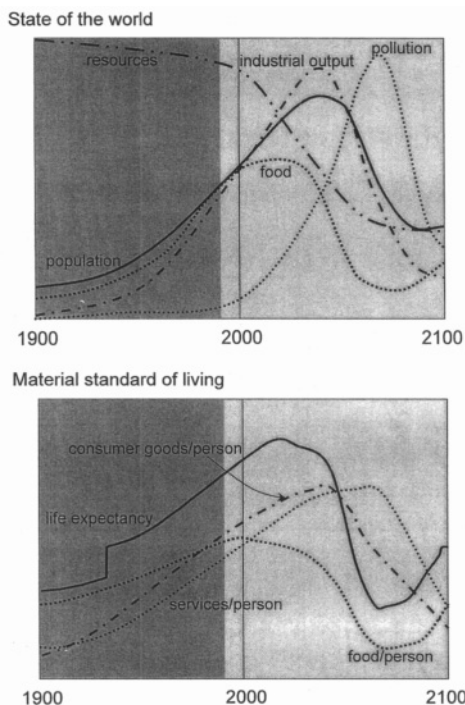
Fig. 15.2. If one is rather liberal in the assumptions as to the natural responses available, industry can grow 20 years longer. Population will increase to more than 9 billion in 2040. These increased levels will generate much more pollution, which will reduce land yield and force much greater investment in fertilizers. Eventually, declining food supplies will raise the population death rate. (Reprinted with permission from D. H. Meadows, D. L. Meadows, and J. Randers, *Beyond the Limits*, p. 135, Chelsea Green Publishing, 1992.)

By the mid-1980s, the general situation was clear enough. The continuation of our present industrial civilization with its reliance on fossil fuels, extensive use of pesticides and fertilizers, and the resulting contaminants will lead eventually to collapse from exhaustion of resources (Meadows, 1972) (Fig. 15.2), and from the health and economic consequences of environmental pollution. There is controversy about the time at which a breakdown would be apparent, but it is clear to those who are conscious of the huge inertia in the industrial processes that are causing the trouble

that waiting until a breakdown is visible before taking remedial action courts the irreversible decay of society.

15.2.   THE ELECTROCHEMICAL ADVANTAGE

It is obvious to those who have comprehended the steadily deteriorating environmental situation that massive changes in our industrial society must be started and completed in the next one to two generations. Some of these changes are purely political in respect to the laws needed. Many changes in technology can be made if there is sufficient will among the populace to see that they are carried out.

There is no doubt that the threat of planetary warming originates in the fact that at present we run our society by utilizing energy obtained from the combustion of coal, oil, and natural gas. Correspondingly, there is no doubt that the major portion of the CO_2 in the atmosphere comes from the way we fuel our transportation system, and that we already have the technology to convert it to one powered by clean electrochemical processes. Photochemical smog is produced from burning gasoline in internal combustion engines and has an electrochemical solution. Acid rain is produced by burning coal and the solution is to use clean alternatives and during the conversion period implement strictly enforced laws to capture the SO_2 still being produced.

In fact, all the insults to our health and well-being that we inflict by breathing and eating in a toxic environment can be mitigated or eliminated either by developing processes to remove the toxins before they reach our food or air, or designing new processes that do not create them.

As this text enters the classroom, electrochemical technology will still be a small part of chemical technology. It is the message of this book that it must become the core of many technological changes needed to ensure a stable and productive future. Within one to two generations, it seems likely that the environmental imperative will force many industries toward electrochemical technology. The reasoning is as follows:

1. It must be a part of our future to move away from combustion of carbon-containing compounds and turn to sources of energy that do not produce planetary warming. These are likely to be renewable resources,[1] which are mostly sporadic and need to be collected in one place and stored and transmitted over long distances. Solar sources are likely to constitute the greater part of these (others are gravitational, wind, waves, and geothermal). They can be coupled with electricity and hydrogen.

---

[1]Although stress is given here to the use of renewable resources to produce energy in place of the polluting fossil fuels, second-generation nuclear energy must not be dismissed from the future energy supply. Our first attempt to use it on a massive scale failed on the basis of the awesome dangers from a plant malfunction (demonstrated by the disaster at Chernobyl in Russia), and on the onerous (as yet unsolved) burden of what to do with its dangerous waste products. However, nuclear energy may make a comeback during the twenty-first century by new technology that is manifestly fail safe and by new scientific discoveries, at present controversial, but which according to the discoverers may make possible nuclear energy with innocuous waste products and in small devices (Nagel, 1997).

2. The direct conversion of the energy of chemical reactions to electricity in fuel cells (Chapter 13) rather than in heat engines will double the energy available and provide clean and relatively simple devices; fuel cells contain no moving parts.

However, the second fuel cell principle $(A+B \rightarrow P + \text{electricity})$ may be applied to industrial processes that now run chemically, producing useless heat. Many of these processes can be run electrochemically, producing the needed product plus transmissionable electricity (Section 13.10). The resulting clean industrial production will then contribute clean electricity to the total energy supply via an energy net that may involve hydrogen for storage, and if the distances involved are long enough, for transmission.

3. Electrochemical processes depend on pressure (concentration) and temperature as do the corresponding chemical ones. However, electrochemical processes also depend on the potential of the working electrode. Thus, the rate at which they take place can be easily controlled, which makes them particularly suitable for pollution control, for example, in purifying factory effluents.

15.3. GLOBAL WARMING

15.3.1. Facts

The rise of atmospheric CO_2 with time is shown in Fig. 15.1. A prediction of the *effects* of this rise is more difficult. Preindustrial CO_2 concentration was 280 ppm (it was about 380 ppm at the century's end). Extrapolation of the present data to 2060 gives a doubling of the preindustrial concentration (560 ppm) and an associated temperature rise of 1.6 K at the equator. The corresponding estimated rise in sea level as a result of polar melting is 50 cm (see Fig. 15.3). Figure 15.4 shows the predicted temperature increases according to various assumptions.

Such estimates are conservative. China, South America, and eventually parts of Africa, are all large regions of the world headed for industrialization with a consequent increase in the rate of CO_2 injection into the atmosphere of as much as 50% over that assumed in the present prediction.

Certain feedbacks will occur, but their effects are difficult to predict reliably. Clouds will increase. Whether they will shield the lower atmosphere by reflecting solar light back into space or increase its heating effect by absorbing more solar light than the ground depends on whether the dominant cloud color is white (reflecting) or black (absorbing). The melting of some arctic ice (the conservatively predicted temperature increase at the poles would be 3.2K) will decrease the reflective power of this region (water is dark, ice is bright) and increase the earth's average temperature. CO_2 is not the only greenhouse gas. As a result of the rising temperature, CH_4 will be released from some tundra, which is abundant in the northern regions of Canada, Alaska, and Siberia. The resulting increased CH_4 would warm the atmosphere further and in turn produce more CH_4, etc.
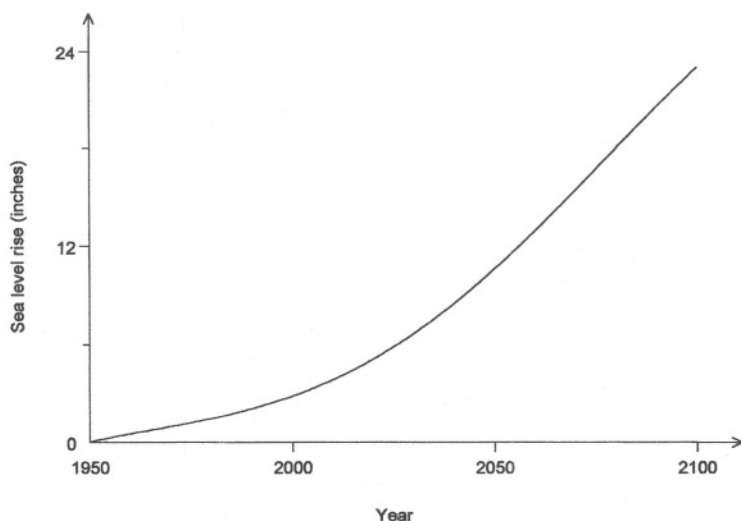
Fig. 15.3. Predicted sea level rise during the twenty-first century (Reprinted with permission of T. Nejat Veziroglu, 1998.)

Apart from these effects, the largest land-based CO_2 sink, the Amazon rain forest, is being destroyed by cattle ranching and lumbering. Deforestation is occurring over the whole land mass, and will lead to an increase in CO_2 as yet unaccounted for.

Although these effects all lead to temperature increases above the minimum quoted above, some effects will act to decrease the estimated rise and should also be taken into account as far as is possible. For example, the increase of smog due to spreading industrialization will lead to a corresponding reduction in the solar light reaching the earth's surface (but how much light will be absorbed and degraded to heat by the low-level smog?).

Estimates of the temperature rise by 2060, taking into account feedbacks, would increase the equatorial temperature by 4.6 °C (Hileman, 1992). However, positive feedbacks may result in an even greater figure. A warming of the tundra that is sufficient to cause a "runaway greenhouse" (positive feedback due to an increase in atmospheric methane as sketched above) would threaten the continuation of life on earth within years rather than decades from the beginning of such releases.

An equatorial temperature increase of 5 °C (9 °F) may not seem too life-threatening. However, this estimate is for the equator. The increase would be 10 °C at the poles (resulting in corresponding melting[2]). A number of predictions of secondary

---

[2]Sufficient warming could cause part of the Antarctic ice cap to become unstable. Should it break away from the land mass, it could cause a rise of several meters in world sea levels.

Meam Global Temperature Increases Depend on Population and Economic Growth

Global Mean Temperature Change, °C
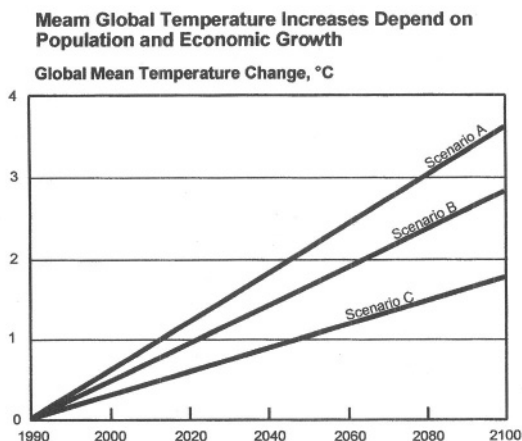
Fig. 15.4. To project global temperature increases during the next century, the Intergovernmental Panel on Climate Change uses scenarios that represent a range of economic, demographic, and policy assumptions. The scenarios include assumptions about the amounts of coal, oil, natural gas, and nuclear fuels that would be consumed, and are used to estimate increases in greenhouse gas concentrations. General circulation models are then employed to calculate likely global temperature changes with these levels of greenhouse gases. A midrange temperature increase of 2.5 °C for a doubling of greenhouse gases is assumed with each scenario.

The base-case scenario B assumes population will increase to 11.3 billion by 2100 and the world economic growth rate will be a moderate 2.3% per year from 1990 to 2100. Scenario A assumes population would also grow to 11.3 billion, but that the world's economy would grow fast, at 3.0% per year. Under scenario C, great constraints would be put on population growth. It would increase from its current 5.3 billion to only 6.4 billion by 2100; annual economic growth would be medium to low at 2.0%, and all deforestation would be halted.

Population is especially important in these scenarios because it is one of the major factors affecting trends in greenhouse gas emissions. The larger the population, the more greenhouse gases are emitted in producing food, energy, clothing, and shelter. (Reprinted with permission from B. Hileman, *Chem. & Eng. News*, April 27, 1992, pp. 7–19. Copyright 1992 American Chemical Society.)

TABLE 15.1
Major Tropical Diseases

Disease	Number of People Now Infected (millions)	Possibility that Climate Change Will Alter Distribution
Malaria	300–500	Extremely likely
Schistosomiasis	200	Very likely
Lymphatic filariasis	117	Likely
Leishmaniasis	12	Likely
Onchocerciasis (river blindness)	17.5	Very likely
American trypanosomiasis (Chagas' disease)	18	Likely
Dengue	10–30 new cases per year	Very likely
Yellow fever	<0.005 new cases per year	Very likely

Source: Reprinted from Richard Monastersky, "Health in the Hot Zone," *Science News*, Vol. 149, April 1996. Copyright 1996 by Science Service. Reprinted with permission from *Science News*.

effects of temperature increases in the United States have been made (Oppenheimer and Boyle, 1992). The temperature changes predicted would unbalance nature. Hurricanes will increase in number and severity. Insect populations in regions at present considered temperate would increase, bringing enhanced disease levels (see Table 15.1). In low-lying countries, the sea rise (more than 1 m by 2060 if the feedbacks are accounted for) could cause significant flooding.

Most nations have accepted proposed limits on the increase of global CO_2 emissions, although they do not add in their protocols that, in the absence of alternative clean energy-producing technologies, this would mean a limit on industrial development and, indeed, on the quality of life. As of 1998, the only major countries that have refused to give a commitment to limit CO_2 emissions are the United States and the United Kingdom. However, without the cooperation of these major powers (together, the two account for about 67% of automotive pollution), the CO_2 burden will continue to increase, largely unchecked.

15.3.2. The Solar-Hydrogen Solution

15.3.2.1. *The Ideas.* The idea that hydrogen (obtained without concomitant CO_2 injection by electrolysis from water) as a medium of energy would be a general solution to environmental problems (Bockris, 1971) was followed by the formation of

an International Association for Hydrogen Energy (Veziroglu,[3] 1974) and the spelling out of a solar-hydrogen scheme (Bockris, 1962, 1975; Justi,[4] 1965).

The basic ideas of a solar-hydrogen energy scheme are first that solar energy should be collected where it is most available, i.e., mostly within 3000 km of the equator. The collection sites would be on useless but highly insolated desert areas or on seaborne platforms (e.g., in the Gulf of Mexico). Major (desert) land areas for solar energy collection are to be found in North Africa, Saudi Arabia, Australia, and southern Spain, among many other places. The electricity thus obtained, at efficiencies of 12–20%, would be used to electrolyze water.[5] It can be shown (Justi, 1982) that storage under varying pressure in pipes would be sufficient to cover the day-night variations in the solar source. The hydrogen from the electrolysis would be pumped up to 100 atm and piped to distant places (from Arizona deserts to Midwestern cities, say) or North Africa to the Ruhr. Here, it would be used in two ways.

To obtain electricity, the hydrogen in the pipes at 100 atm would be used to drive generators, thus achieving an appropriate reduction in pressure of the hydrogen before it is converted to electricity via fuel cells (~65% efficiency). For space heating, the hydrogen would be used directly.

---

[3]T. N. Veziroglu ("Nejat"), a mechanical engineering professor at the University of Miami, is recognized throughout the world as the chief organizer and driver of the idea of hydrogen as a source of energy. In 1974, he was elected president of the International Society for Hydrogen Energy and has remained its president for a quarter of a century. Apart from his remarkable skills as an organizer on an international scale, he has contributed seminally to economic analyses of the interplay between the cost of hydrogen energy from various sources, taking into account the increase in efficiency gained by the use of fuel cells, and the cost of pollution caused by the present fossil fuel system. In this way, he has given powerful support to the concept of a hydrogen economy as a principal solution to planetary warming. Perhaps his best-known paper (with Awad, 1984) is that which began a series of quantitative estimates of the cost of fossil fuel pollution and damage to the environment. With Bockris, he has promulgated the idea of real economics, the net economics of energy after accounting for efficiencies of conversion and pollution costs.

[4]Eduard Justi, a tall man of dignified and even noble appearance (formerly professor of physics at the University of Hannover, Germany), was a prime instigator of the idea of using solar light to drive electrolysis cells to convert water to hydrogen, transmit this clean fuel in pipelines over long distances to cities, and use fuel cells to reconvert the hydrogen to electricity. In 1989 he shared the Augustin–Mouchot prize of the European Solar Energy Society with John Bockris for this suggestion, which he had furthered in books, recording engineering estimates of a detailed kind for trans-European pipelines. Justi's work extended broadly through various researches aimed at creating clean energy and making it available on a very large scale. Although a physicist, he contributed much to electrochemical work, not only in the electrolysis of water from photovoltaic electricity, but particularly in the fuel cell area, where he is known for his 1962 book with Winsel, *Cold Combustion,* and for his seminal work on fuel cell electrodes of uniform porosity, in which a larger fraction of the pores remain active.

[5]It is important to obtain the hydrogen from water (or, e.g., H_2S, see section 15.7.9). However, re-forming of natural gas or methanol (or other fossil fuel) to hydrogen and CO_2 would be a useful beginning toward a clean hydrogen-based world energy scheme. Using the resulting hydrogen in fuel cells would produce energy more efficiently. For this reason, less originating fossil fuel would have to be re-formed than in the present system.
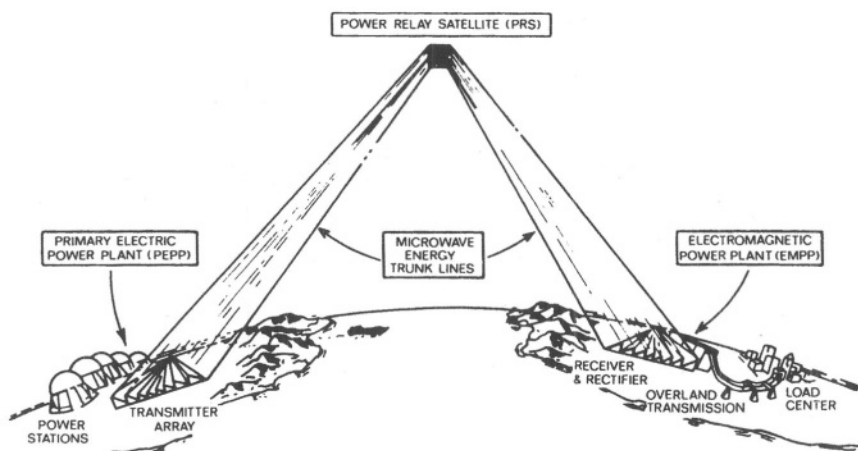
Fig. 15.5. Power relay satellite concept. (Reprinted from J. O'M. Bockris, *Energy, The Solar-Hydrogen Alternative*, p. 175, John Wiley & Sons, 1975.)

Transoceanic transport of the hydrogen would be done according to present schemes after liquefaction, in barges, or after conversion to methanol using atmosphere CO_2, in that liquid. An alternative scheme for very long distance transmission of solar energy (Kraft-Erika, 1974) would be to convert the dc electricity obtained photovoltaically to a megacycle ac frequency and then beam this to a satellite in stationary orbit over the equator. The energy (collected, e.g., from Australian solar-collecting farms) could then be further beamed to distant countries where, on receipt, it would be rectified and used for electrolysis, hydrogen production, storage, and transmission (see Fig. 15.5).

Figure 15.6 summarizes an electrolytically based solar-hydrogen scheme. The scheme may also involve the extraction of CO_2 from the atmosphere and the synthesis of methanol from hydrogen. When the methanol is re-formed in vehicles to give H_2 for fuel cell electricity, the CO_2 evolved would balance that absorbed. No fossil fuels would be involved.

It follows that this scheme—which would lead to a steady-state, high-technology economy without planetary warming or pollution—depends centrally upon electrochemical technology (water electrolysis, fuel cell conversion). If a clean, safe nuclear source were developed that could compete economically with cheap photovoltaics[6],

---

[6]There are several other ways of converting solar energy to electricity. Among them, the cheapest appears to be ocean thermal energy conversion (OTEC). In this technology, the temperature difference between the surface of tropical seas (>25 °C) and that at depth (~ 4 °C) is made to drive turbines by using heat from the sea to boil a suitable liquid, thereby producing electricity. This fluid (e.g., NH_3) would be then condensed by using the 4 °C water pumped up from the deep sea and then recycled as "hot" gas through the turbines.
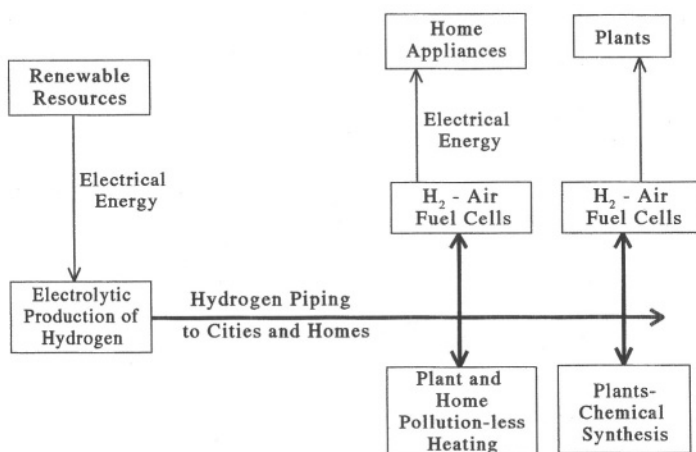
Fig. 15.6. A hydrogen economy. (Reprinted from J. O'M. Bockris, *Energy Options*, p. 259, John Wiley & Sons, 1980.)

it would work continuously and need a storage medium. A hydrogen distribution system would provide such a system.

15.3.3. The Electrochemistry of Water Splitting

Splitting water to obtain hydrogen and oxygen needs an input of energy. Thus, ΔH at 150 °C is +243.03 kJ mol^{-1}, and $\Delta G°$ is +221.49 kJ mol^{-1} at the same temperature. $\Delta S°$ is +51.37 kJ^{-1} mol^{-1}. Hence, $\Delta G°$ becomes less positive with an increase in temperature. At 2000 °C, about 1% of H_2 is in equilibrium with H_2O at 1 atm. These facts have their electrochemical equivalents:

$$E_{cell} = \frac{\Delta G_0}{2F} = 1.229 \text{ V} \tag{15.1}$$

$$\frac{\partial E}{\partial T} = -0.84 \text{ mV } °C^{-1} \tag{15.2}$$

Hence the basic (reversible) thermodynamic standard potential (E_{therm}) of decomposition (as also the overpotentials) decreases as the temperature increases.

For a practical electrolyzer,

$$E_{cell} = E_{therm} + IR + \Sigma \eta \tag{15.3}$$

At 25 °C, E_{therm} is 1.23 V. *IR,* the ohmic potential difference between anode and cathode, can be reduced by good engineering to ~0.1 V (largely a potential difference through a membrane separator) (see also Section 15.3.4). However, this value implies a maximum current density of about 1 A cm^{-2} because at substantially greater current densities, bubbles from the evolving H_2 and O_2 begin to offer significant contribution to the IR losses.[7]

The overpotential terms are those for which there is a frontier in fundamental research on water electrolysis. Electrocatalysis (Sec. 15.3.3) is the field concerned. Older electrolyzers in big plants in the 1990s retained technology that has not yet made use of the possibilities for improvement that arise from the electrocatalytic advances of the past two decades (Murphy and Gutmann, 1984). The lowering of the overpotential at the oxygen anode is vital for the high current densities needed for economic performance. However, the advantage obtained by using noble metals to reduce η_{O_2} is offset by the cost of such materials.[8] A schematic of a modern electrolyzer is shown in Fig. 15.7.

Three possible avenues are available to lower the cost of reducing hydrogen in the large-scale electrolysis of water:

1. Use of steam and temperatures greater than 1000 °C. The advantage of this method lies in the substantial reduction of the overpotential at (say) 1 A cm^{-2} and is accompanied by a significant reduction in the reversible potential as the temperature increases (Pound et al., 1980) (Fig. 15.8).

Temperatures up to 1500 °C reduce the reversible thermodynamic potential for water decomposition from a room temperature value of 1.23 V to ~0.7 V (43%). The cost of electrolytic hydrogen varies linearly with the potential of the cell at the current density being used, since cost of the electricity is the dominating item in the cost of electrolytic hydrogen, high-temperature steam electrolysis would greatly improve the economics. Heat is needed to maintain the temperature of the system, but heat costs only a third of the cost of electricity. So far, very high temperature cells are research items, but 1000 °C cells have been developed in Europe under the nickname "Hot Elly."

2. Pulse electrolysis. It is a fact of electrode kinetics involving gas evolution that pulsing reduces the overpotential for a given current density (Ghorogchian, 1985). With pulses of about 1.5 V, 10^{-3} s in duration, the saving in electricity costs would be about 15%.

---

[7] Among several engineering ideas for reducing IR is one in which the solution is caused to flow so quickly between the electrodes that the H_2 and O_2 bubbles arising from cathode and anode, respectively, have no time to mix and are parted by a separator outside the electrode area where its resistance has no effect on *E,* the total potential. No membrane between the electrodes (with its consequent contribution to the IR loss) would then be needed.

[8] Much progress in reducing the actual amounts of noble metals needed in fuel cells per unit area was made in the 1990s. (Srinivasan, 1993) and it is probable that some of this technology could be used in water electrolyzers that perform reactions reverse to that in H_2-O_2 fuel cells.
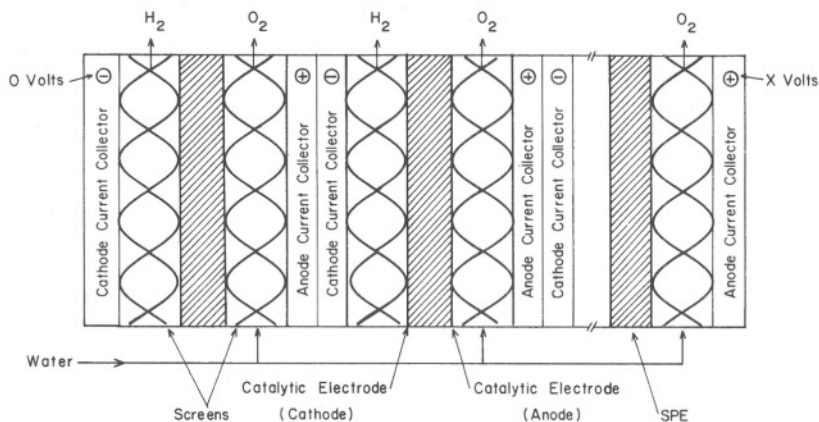
Fig. 15.7. Design of a multicell for a solid polymer electrolyte water electrolyzer. (Reprinted from B. V. Tilak et al., "Electrolytic Production of Hydrogen," in *Comprehensive Treatise of Electrochemistry*, J. O'M. Bockris, B. E. Conway, E. Yeager, and R. White, eds., Vol. 2, Fig. 6, p. 9, Plenum, 1981.)

3. One can electrolyze an aqueous solution in which the anodic reaction is no longer O_2 evolution with its high overpotential. Thus, SO_2, NO, Br^-, Cl^-, and complex organic compounds from slurried coals provide possible anode reactants that reduce the working cell potential, excluding IR, from about 1.7 at 1 A cm^{-2} to various potentials from 1.0 to 1.3 V. At a working potential of 1.15, the saving in electricity costs compared with those of straight water electrolysis below 100 °C would be a significant 36%. Such electrolyzers need commercialization.

At the century's end, the plans of the leading members of the automotive industry were to use hydrogen-fueled fuel cells to run cars electrically. This concept would eliminate the gross pollution (CO, HC, NO, CO_2) from present cars running on the internal combustion of fossil fuels. But there is still a net use of the fossil resources because the methanol would be synthesized from hydrocarbons. Correspondingly, there would still be a net injection of CO_2 into the atmosphere although less than now because of the greater efficiency of fuel cells over combustion engines. A radical change will occur when the carbon needed for the methanol originates in atmospheric CO_2. Then, hydrogen from solar splitting of water would be used to convert the CO_2 to CH_3OH (which becomes a convenient storage medium for the hydrogen from water). When the CH_3OH is re-formed in a car, CO_2 is produced and injected into the atmosphere, but there is no *net* addition because the CO_2 originated from the atmosphere. However, the use of methanol as a storage medium depends on *water* as the source of hydrogen, so that energy scenarios involving methanol storage remain a part of the solar-hydrogen economy.
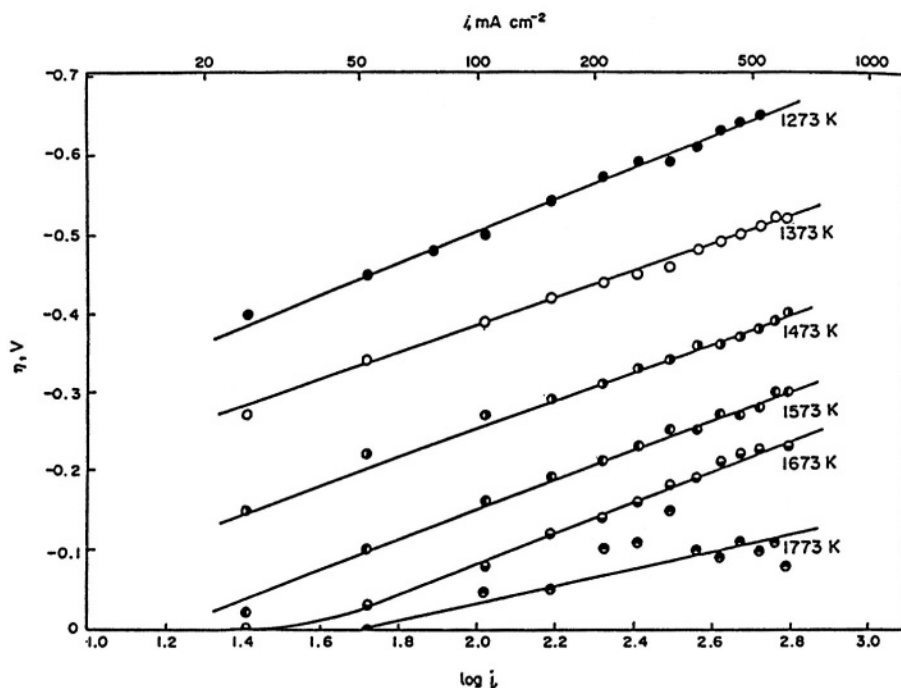
Fig. 15.8. Dependence of cathodic overpotential (η) on log i. (Reprinted from B. G. Pound, D. J. M. Bevan and J. O'M. Bockris, *Int. J. Hydrogen Energy* **6**: 473, 1980. Reproduced with permission of the International Association for Hydrogen Energy.)

15.3.4. The Electrolysis of Sea Water

In a hydrogen-based economy without CO_2 evolution from re-forming, electrochemical water splitting will have to be carried out on a large scale, and the use of highly conducting sea water as the electrolyte merits consideration. Sea water is a more complex liquid than the corresponding 0.4 M solution of sodium chloride. It supports living creatures by means of its O_2 and plankton content, and a system of carbonates and bicarbonates arises from contact with atmospheric CO_2.

Although a description of the technology of sea water electrolysis is beyond our scope, two major aspects of it may be briefly mentioned.

1. The exchange current density for the evolution of Cl_2 on dimensionally stable anodes (TiO_2 with RuO_2 essentially) is about 10^{-5} A cm^{-2} compared with that of O_2 on NiO_2 of about 10^{-10} A cm^{-2}. This would cause (cf. $\approx 2RT/F \ln i/i_0$) a lowering of the anode overpotential at 1 A cm^{-2} of about 0.6 V. The corresponding drop in cell volts (assuming 1.7 V for the hydrogen-oxygen cell) would be about 35%.

2. At the large scale envisaged, the Cl_2 by-product would exceed the market for Cl_2. There are two clean disposal possibilities.

a. Use of the Decon process, which is:

$$H_2O + Cl_2 \xrightarrow{600\,°C} 2\,HCl + 1/2\,O_2$$
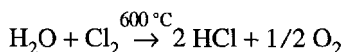
with rejection of the HCl into the sea.

b. Rejection of chlorine into the sea to such a depth that no gas will reach the surface. Chlorine dissolves easily in water.

Alternatively, the electrolysis of brine (sea water concentrated by evaporation) might be advantageous. It would be possible to control the anode potential for the production of $Br_2(l)$ $(E_0 = 1.08\,V)$ to a value < 1.3 V so that the anode product is liquid bromine, not chlorine gas. The liquid Br_2 for which there is no market would be rejected into the sea. Because the Br^- concentration is less than that of Cl^-, measures to accommodate the cathode current density and avoid exceeding the limiting current for Br_2 evolution would be necessary.

15.3.5. Superelectrolyzers

A number of proposals have been made for radical changes in the functioning of water electrolyzers so that they can yield hydrogen at a rate more than 10 times per unit area than that obtained at present (Murphy and Gutmann, 1984). They include:

1. The use of homopolar generators to provide low-voltage, high-current electric power. Such devices would involve a metallic disk spinning between the poles of a supermagnet (10-tesla field strength without the use of liquid H_2 by the use of high-temperature superconductors). Mechanical energy to power the spinning disk can come from hydro or wind resources.

2. Use of brine as the electrolyte for the reasons described in Section 15.3.5.

3. A dimensionally stable anode consisting of light-sensitive lanthanum chromite-titanium dioxide) to encourage the photoevolution of Br_2 (Guruswamy and Bockris, 1979). Light pipes would deliver solar radiation to the anode, thus decreasing the cell potential to significantly less than 1.0 V.

4. Heating the electrodes themselves. This would best be done by utilizing the skin effect, which restricts an ac heating current to the electrode's surface. Increasing the temperature to 100 °C would further reduce the reversible cell potential by about 0.2 V, with a further 0.2 V reduction due to a net overpotential reduction. There may be possibilities for pulse heating for times too short to provoke boiling of the solution in contact with the electrode, but at temperatures far above 100 °C, thus leading to further reduction of the cell potential.

5. Reducing the IR drop through the solution (membrane separator removed) by rapid flow between the electrodes, with the separation of the cathode (H_2) and anode (Br_2) products outside the electrode areas (IR loss reduced to <0.1).

The aim of these concepts is a 1.0 V[9] electrolyzer working at 10 A cm^{-2}.[10]

15.3.6. Photoelectrochemical Splitting of Water

Photoelectrochemical splitting was discussed extensively in Chapter 10. The key point is the use of trace electrocatalysts added to the surface of both photocathode and photoanode to the appropriate extent (Kainthla, Zelenay, and Bockris, 1987; Turner, 1998). If the electrolyzer is to be entirely solar driven, both electrodes must be irradiated. It is difficult to find photoanodes with the appropriate properties. Most of them dissolve electrochemically if used as anodes for O_2 evolution. This can, however, be prevented by using transparent films of nonreactive oxides (Bockris and Uosaki, 1977).

Figure 15.9 shows depth profiles of a Pt 4f and 2s photoelectron signal of a Pt layer on Si and concentration profiles of Pt, Si, and O. Figure 15.10 is a scanning electron microscope picture of a Pt catalyst on Si (Maier, 1996).

Whether it will be better to use photoelectrochemical rather than photovoltaic methods for solar splitting of water needs detailed economic analysis. The photovoltaic conversion of solar light to electricity gives higher efficiencies than have been reached so far in photoelectrochemical devices. However, the latter approach would provide the possibility of producing hydrogen from water and light at a single plant, while in the photovoltaic approach, a separate electrolyzer plant as well as the solar light-collecting farm and equipment would be needed.

15.4. LARGE-SCALE SOLAR-HYDROGEN PRODUCTION

15.4.1. Solar-Hydrogen Farms

Experimental solar-hydrogen plants (Bockris, 1975) using the electrochemical approach to the splitting of water began to be built in France in 1988. Some half dozen plants in various countries are current. However, only two plants supply more than

---

[9]The reversible potential of $Br_2 + 2e \rightarrow 2\,Br^-$ is 1.08 V atm. The pH of sea water is near 7 and 25 °C; hence the reversible potential of $2H^+ + 2e \rightarrow H_2$ is about –0.42 V. The reversible cell potential would be then about 1.5 V. The idea here is to reduce the potential below the thermodynamically reversible value by photoillumination of the anode (photoassisted water decomposition; Szklarczyk, 1983).

[10]The possibility of this very high current density with bromide ions as the anodic reactant requires studies not yet made. The concentration of Br^- in sea water is about 1% that of Cl^-. The anode could be 10^3 times larger in active areas than the cathode. Ultrasonic irradiation of the anode can lead to an increase in limiting current of about 10 times and the use of pulsed high temperatures seems to show (in terms of calculation) the possibility of the projected current density.
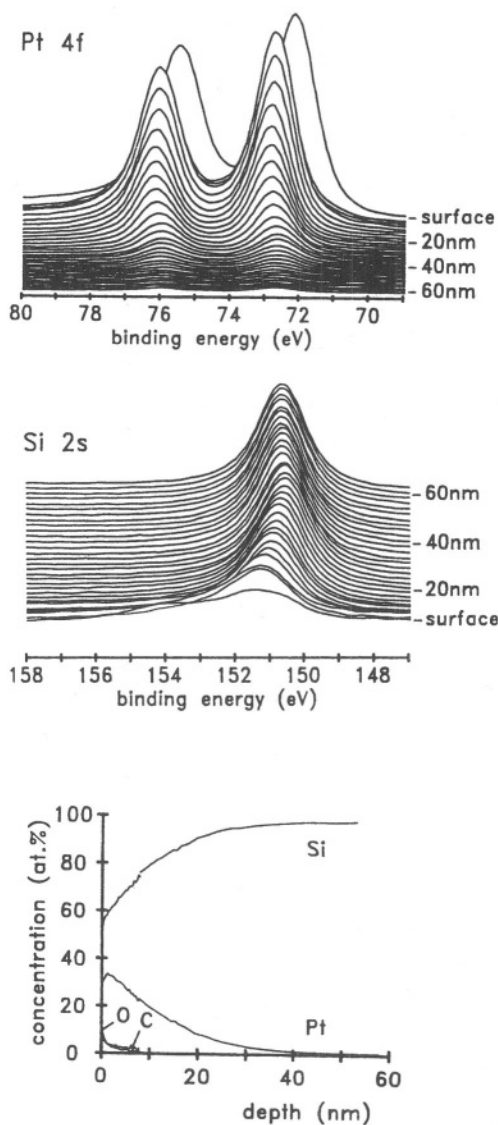
Fig. 15.9. Depth profiles of the Pt 4f and 2s photoelectron signal of the electron beam-evaporated Pt layer on Si and the concentration profiles of Pt, Si, O and C. (Reprinted from C. U. Maier, "Hydrogen Evolution on Platinum-Coated p-Silicon Photocathodes," *Int. J. Hydrogen Energy* **21**: 840, 1996. Reproduced with permission of the International Association for Hydrogen Energy.)
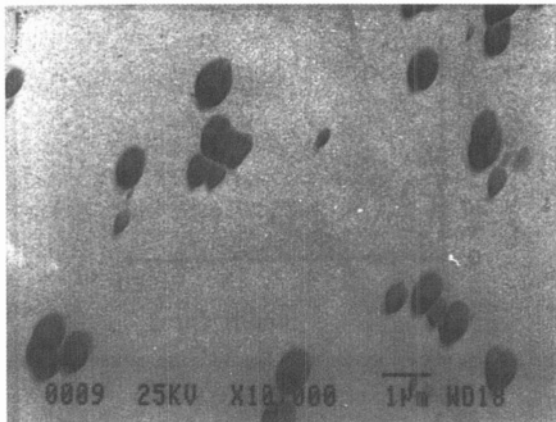
Fig. 15.10. Scanning electron microscope picture of a photoassisted electrochemically Pt-deposited *p*-Si specimen at the surface. (Reprinted from C. U. Maier, "Hydrogen Evolution on Platinum-Coated p-Silicon Photocathodes," *Int. J. Hydrogen Energy* **21**: 836, 1996. Reproduced with permission of the International Association for Hydrogen Energy.)
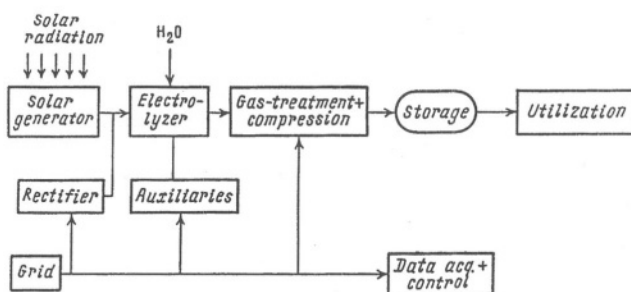
Fig. 15.11. Block diagram of the 350-kW solar-hydrogen plant at Riyadh (Saudi Arabia). (Reprinted from Yu. I. Kharkats, "Electrochemical Storage of Solar Energy," in *Environmental Oriented Electrochemistry*, C. A. C. Sequeira, ed., Fig. 3, p. 473, copyright 1994. Reproduced with kind permission of Elsevier Science-NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.)

100 kW in power. The first is the HYSOLAR plant in Saudi Arabia at 350 kW. A block diagram is shown in Fig. 15.11. The output during a typical day is shown in Fig. 15.12.

The corresponding German plant at Neunburg vorm Wald in Southern Germany (Fig. 15.13) is set to produce 500 kW and is likely by the year 2000 to be the demonstration plant from which the most economic data on the production of solar hydrogen will have been obtained. The plant is still small and its production is less than 0.1% of that of a commercially operating nuclear power plant.
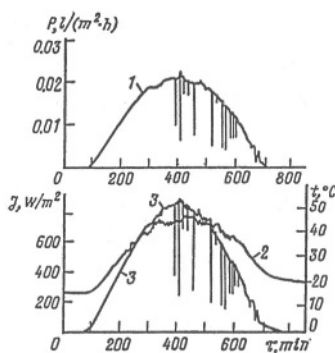
Fig. 15.12. Daily variation in electrolytic hydrogen production rate (1), the solar array temperature (2), and radiation power density (3). Single crystalline silicon solar cells, SPE electrolyzer; location, Cape Canaveral, Florida. The time scale denotes minutes elapsed from 5 a.m. (*Reprinted from* Yu. I. Kharkats, "Electrochemical Storage of Solar Energy," in *Environmental Oriented Electrochemistry*, C. A. C. Sequeira, ed., Fig. 5, p. 477, copyright 1994. Reproduced with kind permission of Elsevier Science-NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.)
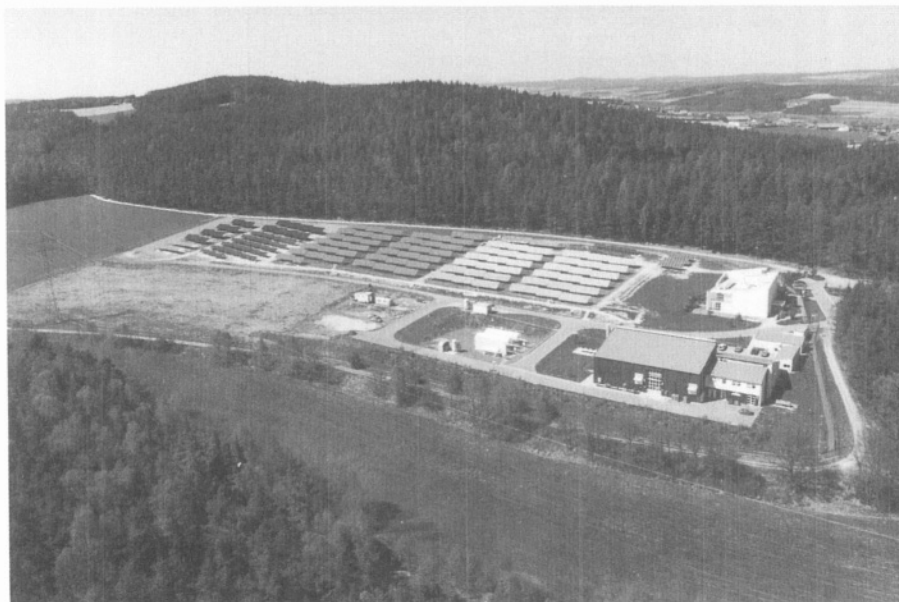
Fig. 15.13. Aerial photo of the overall plant phase I (May 27, 1991). (Reprinted from A. Szyska, "Demonstration Plant, Neunburg vorm Wald, Germany, to Investigate and Test Solar-Hydrogen Technology," *Int. J. Hydrogen Energy* **17**(7), 1992 with permission from A. Szyszka.)

15.5. THE ELECTROCHEMICAL TRANSPORT SYSTEM

15.5.1. Introduction

Environmental pressures in the United States are felt most severely in southern California, particularly from the smogs of San Francisco and Los Angeles. This led in the 1990s to a number of legislative decisions by the California Senate which resulted in state laws on the sale of emission-free cars, with the percentage increasing each year in the new century.[11] Although the starting date of restrictions on the sale of old-type polluting cars has shifted, the importance of the California market and the threat that other states would introduce similar laws, pressured U.S. automakers to take unprecedental action.[12] For example, they obtained (in 1991) an agreement from the federal

---

[11] The decomposition of hydrocarbons in the absence of air gives H_2 and carbon. However, calculation shows that the space needed for storing the carbon powder produced would be too large to provide a practical solution for the mountain of soot, which would grow and grow as the decades went on.

[12] A further influence arose from announcements of proposed electric car manufacture by powerful Japanese and German automakers, with the threat that it would be they who would supply the newly created market.

government to allow collaboration in electric car research and engineering work (otherwise forbidden by an act aimed at reducing monopolies).

At first, a large program of research in new kinds of batteries (Chapter 13) was carried out. However, even the best of the new batteries (e.g., the Li-ion batteries) still involved excessive costs, unacceptable recharging times, and in particular polluting CO_2 injection from the new electricity-producing plants.

In 1992, it was pointed out (Billings) that an immediate fuel cell solution allowing a longer range than that of vehicles powered with internal combustion engines was at hand. It would be a good compromise to make the hydrogen that fuel cells need from carbonaceous fuels through steam re-forming at gas stations or on board the vehicle. Thus, while CO_2 would still be injected into the atmosphere as a consequence of the steam re-forming reaction,

$$CH_4^4 + 2H_2O \rightarrow CO_2 + 4H_2$$
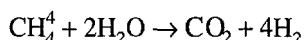
the fact that fuel cells operate at twice the efficiency of internal combustion engines means that upon complete conversion of a country's automotive fleet to (low-cost) fossil-derived H_2, the CO_2 production would be about half of that emitted by internal combustion (Fig. 15.14). Further, most of the other polluting emissions from internal combustion engines (particularly the smog-causing unsaturated hydrocarbons) would be eliminated.[13]

Even these ideas (which have been in the open literature since 1993) did not tempt U.S. automotive makers to switch to fuel cells as the power source for electric cars. The lead was finally taken by Daimler-Benz, in Germany, the first manufacturers of passenger automobiles with internal combustion engines in the world. In 1996 this company demonstrated fuel cell-driven passenger cars with methanol as the originating liquid to be taken on board and re-formed to hydrogen. This fuel is then used in fuel cells to power the cars' electric motors.[14]

Daimler-Benz's announcement was followed by analogous proposals by other automotive companies. Chrysler proposed to use gasoline as the originating fuel (which Daimler-Benz had rejected as being more difficult to re-form to hydrogen than methanol).

These proposals have led to a shift in the paradigm. The infrastructure would be the same as at present. The public would continue to visit gas stations and pour a liquid into their cars in the good old way, and at about half the fuel cost (because of the doubling of efficiency offered by the fuel cell). The oil companies would continue to sell products originating in fossil fuels although less of them (but the degree of necessary refinement will be reduced). By 1997 it was clear that the direction taken by the major automakers in the transformation toward totally clean cars had undergone

---

[13] Emissions of NO, CO, and SO_2 in the fuel cell (steam re-former) scheme would be zero, whereas for battery-powered cars, the electricity-producing plants would emit these pollutants in substantial quantities.

[14] The fuel cell–electric motor combination is termed "the electrochemical engine."

Another Option for Cleaner Driving

Fuel cells offer the potential to power vehicles while causing far less pollution than other types of engines. Here is how they compare, based on the emissions from the engine and producing the fuel to run them.

FUEL-CELL MOTORS

Using hydrogen gas	0 0 0

Total amount of each pollutant emitted for each mile traveled, in milligrams:

Using natural gas	2 52 11

■ Sulfur dioxide

Using methanol	2 39 8

▨ Nitrogen oxides

☐ Carbon monoxide

COMBUSTION ENGINES

Gasoline powered	48 205 700

Diesel powered	48 350 870

BATTERY-POWERED ELECTRIC MOTOR

(Emissions from generation of power to charge battery)	145 143 40

Source: Daimler-Benz

Fig. 15.14. Another option for cleaner driving. (Reprinted from Daimler-Benz, November 1997.)

a 45° turn. Electric cars would cut pollution. But it would be by the use of fuel cells, not batteries,[15] to provide the electric power. The pollutants CO, CH, NO, and SO_2 would be eliminated (Fig. 15.14). CO_2 production would be cut by half.

The initiatives taken by these companies would have been delayed except for the technical advances made by a single company, the importance of which became apparent only in the 1990s: the Ballard Company of Vancouver, British Columbia, and research led by D. P. Wilkinson. German, Japanese, and finally U.S. automakers turned to this company (which had demonstrated a full-sized electrochemically powered bus in 1995) for the development of fuel cells for their cars.

15.5.2. Electrochemically Powered Cars

The main point on which the transfer to a clean electrochemical transportation system has stumbled in the past has been the power source, which, with batteries, could provide only a small range, a long charging time, and continued pollution from the

---

[15] The BMW Company (which encourages an image "more sporty" than that of the rival Mercedes cars of Daimler-Benz) is also active in fuel cell-driven cars. However, their plans involve the use of liquid hydrogen, providing a range of 1000 km.

electricity production plants. With the change to fuel cells, these problems are solved because the range of a fuel cell–electric motor-powered car is potentially greater than that of a car powered by internal combustion of a fossil fuel, and the refueling time is the same (the hydrogen carrier being methanol or gasoline).

Electric cars have several advantages over cars running on a combustion principle. The main ones are silence in operation, absence of vibration, diminished maintenance, and the availability of electromagnetic braking, the electricity from which can be used to produce extra H_2 fuel from the water obtained when H_2 is combined with airborne oxygen.

There are other possibilities with fuel cells in cars. For example, the so-called super flywheel has a higher power density than the fuel cell. Fuel cell-powered electric motors would drive a central flywheel and the energy generated would be transferred in a liquid medium to hydraulic motors on each wheel.[16] Alternatively, the fuel cell would drive a pump to compress air to be used in air turbines on all wheels. Devices of this kind may be in the second generation of electrochemically powered cars. Toward the century's end they were still designs and one-off model cars that use solar collection on the car, on-board hydrogen production to store the solar energy, fuel cells, and a central electric motor to drive a flywheel serving hydraulically operated turbines on each wheel. The claimed acceleration and performance of such a vehicle (Krassner, 1991) is better than that of present sports cars.

Yet another possibility relies upon the remarkably high capacity of the electrical double layer. The so-called condenser batteries (Chapter 13) would provide a supplementary energy supply analogous to that of superchargers on internal combustion vehicles. Extra electricity storage capacity would be available with acceptable energy density for startup, acceleration, and passing (Conway, 1998).

15.5.3. The Fuel Cell

The electrochemistry of fuel cells is discussed in Chapter 13. Superior catalysts for the oxygen cathode are at the frontier of research.

Storage of pure hydrogen on board a vehicle would have the advantage of a super-long travel range (>1000 km). The hydrogen could be stored in high-pressure tanks of lightweight alloys. The BMW company has chosen cryogenic storage. In an interim period before the general conversion to nonfossil energy sources, conversion of a fossil fuel to hydrogen in re-forming stations (rightly called "gas stations") might give the optimum economics. It would give a car taking on board liquid hydrogen the advantage of ranges more than twice those now available with polluting internal combustion motors. Although some NO_x is still produced as a by-product of the re-forming reaction, this could be eliminated by catalytic converters at midsized

---

[16]A car using a central electric motor activating a fluid drive to turbines on all wheels was designed, built, and operated by the engineers in the physical science workshop at Flinders University, Ade laide, Australia, in 1976.
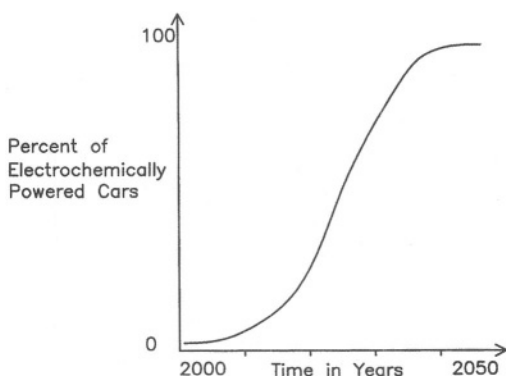
Fig. 15.15. A speculative portrayal of the growth of fuel cell-based electric transportation.

re-forming plants at gas stations, rather than by smaller devices aboard the cars.[17] A speculative estimate of the growth pattern for electrochemical transportation (which can be extended to trains, ships, and planes) is given in Fig. 15.15.

15.6. THE FIXING OF CO_2

15.6.1. Introduction

Some 10^9 tons of CO_2 are added to the atmosphere each year, and during the next half century, if present practices continue, the CO_2 concentration will rise to at least 500 and perhaps even 600 ppm (about a 48% increase over the concentration at the year 2000). One way to deal at least partially with this undesirable development would be to "fix" the CO_2 into, e.g., the simplest practical liquid H_2 carrier, methanol (compare the use of methanol as a hydrogen source for fuel cells in electrochemical transportation and in other industrial machinery).

On rough assumptions concerning the average yearly consumption of gasoline per car, some 5×10^{13} mol of MeOH would be needed to drive the cars of the world on internal combustion; about half that would be needed for cars powered by fuel cells. Taking the CO_2 emission at 2050 as 5×10^9 tons/yr, this is equivalent to about 10^{14} mol of CO_2. Thus, it is not fanciful to conceive of fixing most of the new CO_2 output

---

[17]More advanced possibilities lie beyond those within the 30–50-year horizon of this book. The corresponding application of fuel cells to rail and sea will follow, not only for environmental but also for economic reasons. The possibilities of towers bearing photovoltaic panels and collecting air for CO_2 extraction, may make possible solar conversion to methanol at individual gas stations. Airborne vehicles used as observation platforms collecting solar light in flight and storing it in hydrogen for night use in fuel cells could stay aloft indefinitely.

in a form that would (along with water as the source of hydrogen) provide a convenient hydrogen carrier for automotive fuel in fuel cells.[18]

15.6.2. The Possible Reduction Product

There are some six pathways for CO_2 reduction. At pH 7 and on the normal hydrogen scale of potential, these reactions are as follows:

$$CO_2(g) + 8H^+ + 8e \rightarrow CH_4(g) + 2H_2O \qquad E° = -0.24 \text{ V}$$

$$CO_2(g) + 6H^+ + 63 \rightarrow CH_3OH(aq) + H_2O \quad E° = -0.38 \text{ V}$$

$$CO_2(g) + 4H^+ + 4e \rightarrow HCHO(aq) + H_2O \quad E° = -0.48 \text{ V} \qquad (15.4)$$

$$CO_2(g) + 2H^+ + 2e \rightarrow CO(g) + H_2O \qquad E° = -0.52 \text{ V}$$

$$CO_2(g) + 2H^+ + 2e \rightarrow HCOOH(aq) \qquad E° = -0.61 \text{ V}$$

$$2CO_2(g) + 2H^+ + 2e \rightarrow H_2C_2O_4(aq) \qquad E° = -0.90 \text{ V}$$
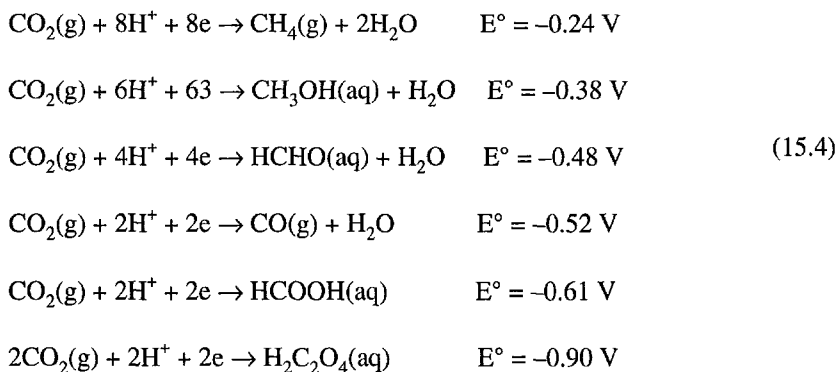
It is clear at once that in aqueous solutions, the reduction of CO_2 will be difficult because of competition with hydrogen evolution, so that catalysts with low i_0 for this latter reaction will be best. Nonaqueous solutions would offer the advantage of a greater CO_2 solubility and no competition with H_2, which must be balanced against reduced conductivity and the difficulty of keeping the system anhydrous over longer times.

15.6.3. Reduction of CO_2 on Metals

Fischer and Prziza (1914) were the first to obtain a near 100% electrochemical reduction of CO_2 (to HCOOH). The major problem is the competition with hydrogen. They achieved success by using amalgamated zinc and high CO_2 pressure together with an overpotential greater than 1.8 V. The current efficiency for CO_2 reduction on various metals is shown in Fig. 15.16. Indium and tin are the best electrodes for the electrochemical reduction of CO_2 in an aqueous solution because each of these metals has a high hydrogen overpotential.

Electrocatalytic possibilities with CO_2 reduction in aqueous solution are surprising. On copper at 0 °C, CH_4 is the main product from electrolysis, and on a molybdenum cathode at room temperature, it is methanol (Hori, 1980). Using lead in a porous gas diffusive electrode, it has been possible to obtain HCOOH at 100 mA cm^{-2} (Hallmann, 1991). Macrocyclic compounds catalyze the reduction of CO_2 to CO to

---

[18]If one laid this burden on half of a world population of 5 billion, it would be equivalent to 5 kg of CO_2 reclaimed from air per day per person.
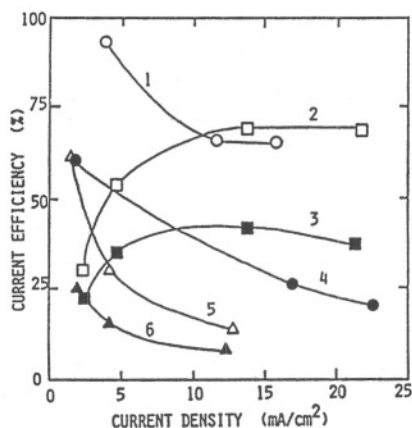
Fig. 15.16. Current efficiencies for reduction of CO_2 to formic acid in an 0.1 M Li_2CO_3 solution at 25 ± 1 °C at various electrodes. 1, In; 2, Sn (previously anodized); 3, Sn; 4, Zn; 5, Pb; and 6, Cd. (Reprinted with permission from I. Taniguchi, "Electrochemical and Photoelectrochemical Reduction of Carbon Dioxide," in *Modern Aspects of Electrochemistry*, J. O'M. Bockris, R. White and B. E. Conway, eds., No. 20, Fig. 1, p. 329, Plenum, 1989.)

various degrees (Enyo, 1991), depending greatly on the modification of a CO-based macrocycle on glassy carbon (Fig. 15.17).

Another approach can be taken to converting CO_2 to methanol, and this is to use now available catalysts for the chemical synthesis:

$$CO_2 + 3H_2 \xrightarrow{\text{cat}} CH_3OH + H_2O$$
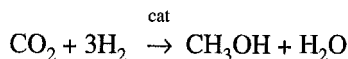
Electrochemistry would play a part in fixing the necessary CO_2. Air would be filtered through hydrophobic, microporous, hollow fiber membranes and dissolved in 1 M KOH. This produces a high surface area of bubbles, which maximizes dissolution. Electrolysis of the resulting K_2CO_3 solution gives CO_2 and H_2. A membrane would be used and the KOH used regenerated in the cathode compartment. Extra hydrogen would be needed for the above catalytic reaction and to make up $3H_2$ for 1 CO_2; this would be formed in parallel in a normal electrolyzer. Together with the H_2 and CO_2 from the carbonate electrolysis, this provides the constituents for the chemical synthe-
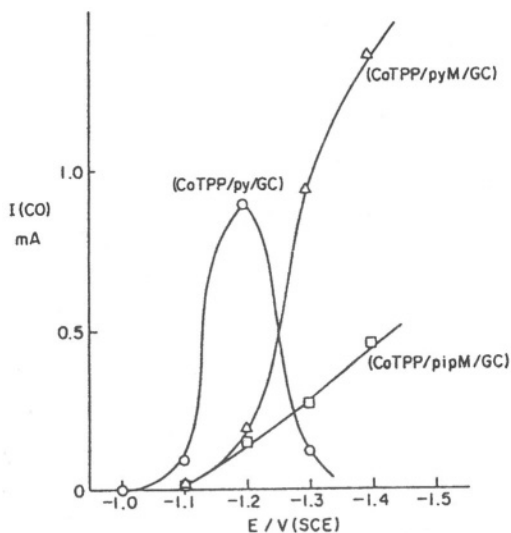
Fig. 15.17. Current in milliamperes cm^{-2} for CO production from CO_2 using CO^{2+} tetraphenyl porphyrins modified with various pyridyl derivatives. (Reprinted with permission from T. Atoguchi, A. Aramata, and M. Engo, "CO_2 Reduction by Macrocyclic Transition Metal Complex-Modified Electrodes," in *Proc. International Symposium on Chemical Fixation of Carbon Dioxide*, p. 338, Fig. 5, 1991.)

sis of CH_3OH [cf. the possibilities available for the production of low-cost H_2 through the use of novel forms of electrolysis in Section 15.4]. The process (Stucki, 1995) could be carried out at individual gas stations (Fig. 15.18).

The advantages of the greater solubility of CO_2 in nonaqueous solution (e.g., acetonitrile) along with the absence of competing H_2 evolution need further development. Glycolic acid is the product of CO_2 reduction in acetonitrile.

15.6.4. The Mechanism of CO_2 Reduction

The classic work on the mechanism of electrolytic CO_2 reduction in aqueous solution was carried out by Park, Anderson, and Eyring but was not published until 1969. They determined the Tafel slopes and reaction orders with respect to CO_2 and found for the latter, zero and for the former two regions, 0.09 in a lower current density and 0.20 in a higher one. They found these results (on Hg) to be consistent with the pathway:
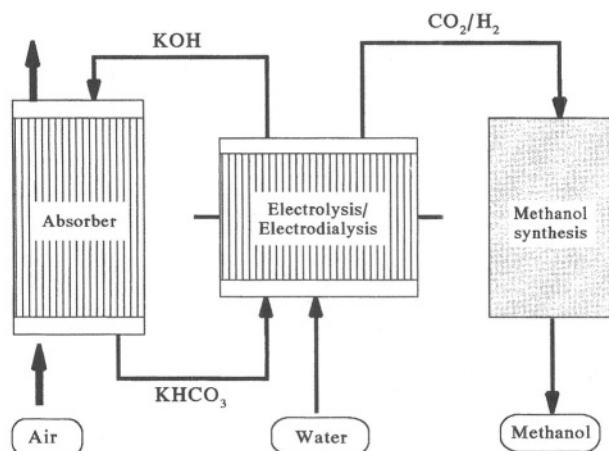
Fig. 15.18. General process scheme for the production of methanol from renewable resources. (Reprinted from S. Stucki, A. Schuler, and M. Constantinescu, "Coupled CO_2 Recovery from the Atmosphere and Water Electrolysis: Feasibility of a New Process for Hydrogen Storage," *Int. J. Hydrogen Energy*, Vol. 20, p. 654, Fig. 1, 1995. Reproduced with permission of the International Association for Hydrogen Energy.)

$$CO_2 + e_0^- \rightarrow CO_{2(ads)}^- \tag{15.5}$$

$$CO_{2(ads)}^- + H_2O \rightarrow HCO_2^- + OH^- \tag{15.6}$$

$$HCO_2^- + e \rightarrow HCO_2^= \tag{15.7}$$
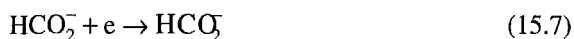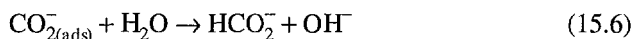
In the lower current density region, reaction (15.7) is rate determining and at the higher current densities, reaction (15.5) is rate determining.

The CO_2^- assumed in this mechanism of the reduction of CO_2 on an Hg electrode was detected through an early application of FTIR (Chandresekaran and Bockris, 1987) to Pt in aqueous acetonitrile; the mechanism was found to have a pathway similar to that in aqueous solutions on Hg and Cu.

$$CO_{2_{soln}} \rightleftharpoons CO_{2_{ads}} \tag{15.8}$$

$$CO_{2_{ads}} + e_0^- \rightarrow CO_2^- \; (\text{rds}) \tag{15.9}$$

$$CO_{2_{ads}}^{-} + H_2O + e_0^{-} \rightarrow HCOO^{-} + OH^{-} \tag{15.10}$$
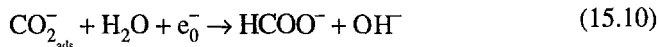
In CO_2-MeOH mixtures on Cu, a current density of 400 mA cm^{-2} was measured (Fujishima, 1995). Depending on the cation present, CO or CH_3COOH were products.

15.6.5. Photoelectrochemical Reduction of CO_2

Nonaqueous solutions have been used exclusively in this process. The first C–C bond to be formed photoelectrochemically from CO_2 was oxalic acid in dimethyl formamide (DMF). A two-photoelectrode cell was used involving a *p*-type GaP electrode and an *n*-type $LaCrO_3$-TiO_2 electrode (Guruswamy, 1979). Photoelectro-catalysis is marked for *p*-type semiconductors in DMF–0.1 *M* tributylammonium phosphate containing 5% water (Taniguchi, 1986) (Fig. 15.19). CdTe was the best catalyst in reactions that formed mainly CO.
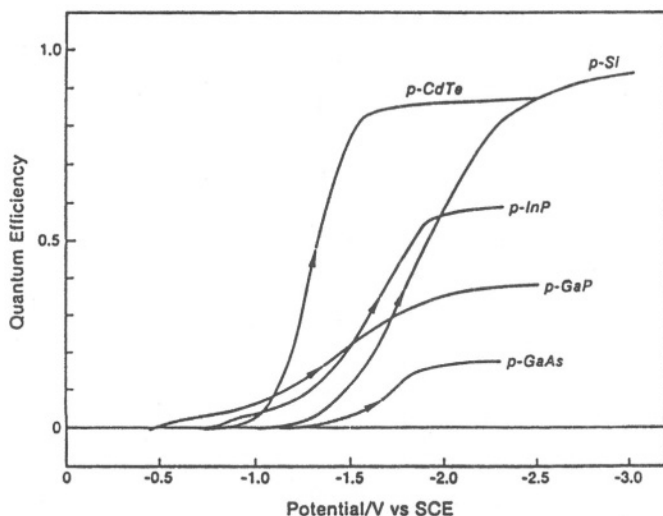
Fig. 15.19. Quantum efficiency versus potential at various *p*-type semiconductors in a DMF-0.1 *M* TBAP solution containing 5% water under a CO_2 atmosphere. Monochromatic light of 600 nm was used for *p*-Si, *p*-InP, *p*-GaAs, and *p*-CdTe, while light of 400 nm was used for *p*-GaP. Scan rate: 0.1 V/s. (Reprinted from I. Taniguchi, "Electrochemical and Photoelectrochemical Reduction of Carbon Dioxide," in *Modern Aspects of Electrochemistry*, J. O'M. Bockris, R. White, and B. E. Conway, eds., No. 20, Fig. 7, p. 357, Plenum, 1989.)

A remarkable example of photoelectrocatalysis was studied by Bockris and Wass (1989). They used 5% aqueous DMF solution with p-Si and p-CdTe as the photoelectrodes. Photoelectrocatalysis was observed by decorating the electrode surfaces with a number of different metal atoms. However, a 10^3 times acceleration of the rate of the photoreduction of CdTe was found in the presence of 18 crown 6 ether[19] in dimethyl formamide solution (Fig. 15.20). The presence of both the NR_4^+ ion and the appropriate crown ether was the key to the acceleration of the reaction rate. The suggested sequence is:

$$NR_4^+ + e \rightleftharpoons NR_{4ads} \tag{15.11}$$

$$NR_{4ads} + CO_{2ads} \rightleftharpoons NR_4^+ + CO_{2ads}^- \tag{15.12}$$

$$CO_{2ads}^- + H^+ + e \rightarrow CO_{ads} + OH^- \quad \text{(rds)} \tag{15.13}$$

$$OH^- + H^+ \rightarrow H_2O \tag{15.14}$$
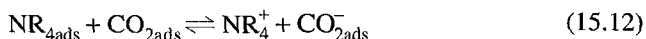
Equation (15.11) gives the coverage, θ_{NR_4} as:

$$\theta_{NR_4} = \frac{K_1 a_{NR_4^+} \exp(VF/RT)}{1 + K_1 a_{NR_4^+} \exp(VF/RT)} \tag{15.15}$$

Equation (15.16) gives

$$\theta_{CO_2^-} = \frac{K_1 K_2 p_{CO_2}}{a_{NR_4}} \frac{a_{NR_4^+} \exp(VF/RT)}{1 + K_1 a_{NR_4^+} \exp(VF/RT)} \tag{15.16}$$

$$v = k_1 a_{H^+} K_1 K_2 \, p_{CO_2} \frac{\exp(VF/RT)}{1 + K_1 a_{NR_4^+} \exp(VF/RT)}$$

$$\times \left[1 - \frac{K_1 a_{NR_4^+} \exp(VF/RT)}{1 + k_1 a_{NR_4^+} \exp(VF/RT)} \right] \exp(-\beta VF/RT) \tag{15.17}$$

If the CdTe surface is homogeneous, then, with $K_1 a_{NR_4^+} \exp(VF/RT) \gg 1$,

---

[19]Crown ethers are cyclic compounds of ethylene glycol containing the sequence $-O-CH_2-CH_2-O-CH_2-CH_2-O-$ etc. The 18 refers to the total number of atoms in the ring and the 6 to the O's.
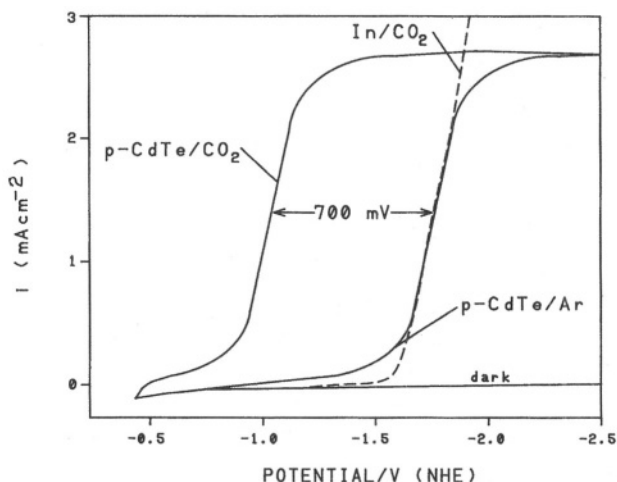
Fig. 15.20. Current–potential curves at a *p*–CdTe electrode in DMF-0.1 *M* TEAP/5% H_2O. (Reprinted from J. O'M. Bockris and J. C. Wass, *J. Electrochem. Soc.* **136**(9): 2522, Fig. 1, 1989. Reproduced with permission of The Electrochemical Society, Inc.)

$$v = \frac{k_1 a_{H^+} K_2 \, p_{CO_2}}{a_{NR_4^+}} \exp(-3VF/2RT) \tag{15.18}$$

This mechanism leads to a Tafel slope of $2RT/3F$, which is inconsistent with the measured $2RT/F$. However, it can be shown that it gives the experimentally observed slope if the CdTe surface is heterogeneous (as observed) and contains nonphotoactive (A) sites on which the NR_4^+ adsorbs and produces CO_2^- (as in the first part of the above mechanism), while the B sites are photoactive and involve the rest of the mechanism shown above. The effect of crown ethers that fit the size of the NR_4^+ (Fig. 15.21) ions is to enhance their specific adsorption and hence the reaction rate.

15.6.6. Conversion of an Organic Compound in Photoelectrochemical Fixing

It is usually thought that the fixing of CO_2 has as its aim the formation of a usable compound, particularly methanol. However, it is possible to use atmospheric CO_2 in the electrosynthesis of organic compounds, and this is well exemplified by the conversion of benzyl chloride to phenyl acetic acid. This was carried out (Uosaki and Nakabayashi, 1993) in a cell in which the electrolyte was acetonitrile and DMF. The photoelectrodes were GaAs, InP, and GaP; the last was the most promising. An Mg
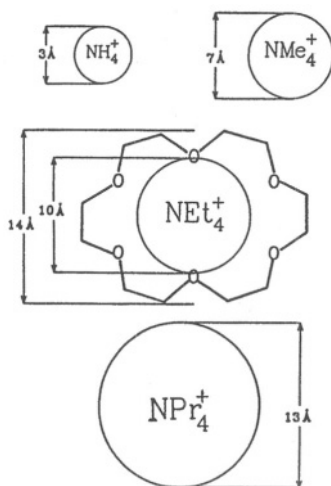
Fig. 15.21. Molecular drawing depicting the size of 18-crown-6, with the tetraethylammonium cation fitting into the central cavity. (Reprinted from J. O'M. Bockris and J. C. Wass, *J. Electrochem. Soc.* **136**(9): 2526, Fig. 12. Reproduced with permission of The Electrochemical Society, Inc.)

anode helped the cell to be stand-alone under solar irradiation (Fig. 15.22). The interpretation of the i–V plot leads to the conclusion that the electrode is functioning as a Helmholtz (and not a Schottky) semiconductor, i.e., it has a high degree of surface states (Section 10.5.2).

15.6.7. Prospects in the Electrochemical Reduction of CO_2

A consensus might be found in support of the statement that the photosynthetic reaction:

$$nCO_2 + nH_2O \xrightarrow{h\nu} (CH_2O)_n + nO_2$$

---
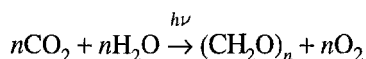
[20]Thus, it forms the green plants, grass, and trees. These plants are ingested by animals, some of which are consumed by humans. In this sense solar energy may be said to drive nature. Szent-Gyorgyi (who pioneered the idea of regarding some bioreactions in terms of semiconductor electrochemistry) was the first (1978) to point to green plants as hydrogen storers of solar energy.

is the most important reaction in nature.[20] Its mechanism has two parts. The first involves the photoelectrochemical formation of H and O_2.

One might, then, see one aim of a photoelectrochemical technology as the trapping of solar light onto two-electrode photoelectrochemical cells to convert CO_2 to CH_3OH. The previous section shows the remarkable possibilities of the first decade of intensive work on the fixing of CO_2. The development of technologies that could contribute to the solution of planetary warming by this route seems to be largely determined by research funding.

A reasonable goal of electrochemical science in the twenty-first century is the large-scale fixing of atmospheric CO_2 by photoelectrochemical reaction to form MeOH. A still further goal would be based upon synthetic food and textile production from water (hence H_2 by means of electrolysis), atmospheric CO_2, and N_2. The basic materials are all available from the constituents of the atmosphere, bacteria, solar light, and the electricity obtained from them.
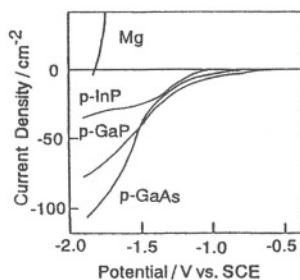
Fig. 15.22. Electrochemical behaviors of *p*-type photoexcited semiconductors and Mg electrodes in CO_2-saturated acetonitrile solution with 50 m*M* of benzyl chloride. (Reprinted from H. Ueda, K. Nakabayashi, Z. Ushizaki, and K. Uosaki, "A Photoelectrochemical Fixation of Carbon Dioxide. Spontaneous Up Quality Conversion of Organic Compound," *Chem. Lett.* **190**: 1748, Fig. 2, 1993. Reproduced with permission of The Chemical Society of Japan.)

15.7. REMOVAL OF WASTES

15.7.1. Introduction

When the question of environmental pollution comes up, the focus is first on the effects of the combustion of fossil fuels used in transportation and the SO_2 emitted from coal-burning electricity plants. However, although these important matters fill about half of the picture of environmental damage, there are many other areas of concern. An early discussion on how electrochemical technology could solve some of these other problems was published in 1972 under the title *The Electrochemistry of Cleaner Environments* (J. O'M. Bockris, ed.).

In general, until the 1970s, most municipal authorities were satisfied with using incineration to dispose of many industrial and household wastes. It gradually came to public consciousness, however, that incineration simply spread the products of incomplete combustion over the surrounding area. Not only were some of these products highly toxic (e.g., chlorobenzene), there was also the daunting possibility that some radioactive wastes could be volatilized as $PuO_2(OH)_2$ and $UO_2(OH)_2$.[21] The realization that, however high the incinerator chimney, its products would land somewhere, led to a change in attitude. Since the early 1990s, there has been a trend in technologically active countries to develop alternative electrochemical methods of dealing with wastes. The obvious advantage of an ambient-temperature aqueous process is that it does not spread waste products far and wide and it is possible to control the potential and thus the degree to which oxidation should be driven to completeness, i.e., to CO_2. Three kinds of electrochemical processes are involved:

1. Anodic oxidation using a large-area inert electrode.
2. Mediator processes. The oxidation occurs homogeneously using such ions as Fe^{3+} and Ag^{2+}. The electrodics involve the reoxidation of the corresponding Fe^{2+} and Ag^+.
3. Electrogenerative processes. Processes can be devised in which some wastes act as a fuel for a fuel cell, the corresponding cathodic process being the electroreduction of oxygen. Electricity is the by-product. Contaminated kerosene and industrial alcohol wastes would lend themselves profitably to such a pathway.

The following section describes a few examples of the electrochemical treatment of wastes.

---

[21]Evidence gathered in 1997 at the Hanford facility for nuclear waste storage shows some leaking into the surrounding soil. On the other hand, the most potent nuclear waste products (plutonium, cesium, and strontium) do not move far (they become adsorbed on soil), although technetium (with a half-life of 250,000 years) has reached the water table.

15.7.2. Waste Water

The range of impurities in so-called waste water is 5–500 ppm. The water becomes acceptable for many purposes if the level is reduced to 500 ppb. However, the impurities are largely nonconducting organics, not ions, so that from the electrochemical point of view, the difficulty is the high resistance. There are two key elements in the process (Hitchens, 1991):

1. The electrode-catalyst (Pt-Ir) is chosen so that i_0 for the O_2 reaction is particularly small and therefore oxygen evolution will be suppressed and most of the current will go to the oxidation of the organics.
2. The electrode material fills as much as half the solution volume so that the current path is short and the wasteful IR drop tolerably small.

As an example,

$$HCOOH + H_2O \rightarrow CO_2 + \tfrac{1}{2} O_2 + 4H^+ + 4e_0^-$$
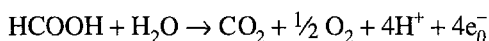
The protons produced in the oxidizing reaction pass through the proton exchange membrane and give rise to the evolution of H_2 on a cathode.

15.7.3. Sulfur Dioxide

The combustion of coal and oil to make electricity produces SO_2, which reaches the atmosphere and is the primary cause of acid rain. Dissolved SO_2 could instead be used to make electricity in a fuel cell process:

$$2SO_2 + 2H_2O \rightarrow 2SO_3 + 4H^+ + 4e_0^-$$

$$O_2 + 4H^+ + 4e_0^- \rightarrow 2H_2O$$
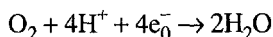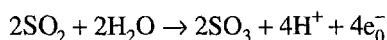
At present, SO_2 is removed from stack gases by a chemical reaction with CaO to form $CaSO_4$. However, the disposal of the solid sulfate presents a logistic and economic problem.[22] Should it be transferred over long distances to take the place of the mined coal?

The extraction of metals from sulfide ores is still carried out in some plants by chemical oxidation, with the result that great quantities of SO_2 are ejected into the atmosphere. In the environmentally friendly electrochemical process, the sulfide ores are crushed to a pulp and the suspension oxidized by hypochlorate in solution to form

---

[22]The final product of the fuel cell process would be sulfuric acid. If the amount produced exceeded the market for this product, its proper disposal in public waters could be done in such a way that the pH change would be reduced to negligible proportions by dilution.

MSO_4 from which the M (M = Hg, Ag, Mo, Cu, and Sb), can be selectively electrodeposited.

15.7.4. Removal of Metals: Aquifers

Industrial effluents from plating baths, some factory waste water, and wastes from photographic developer units all contain an amount of metal ions worth recovering. The solutions are passed through a packed bed consisting of, for example, lead shot, with a certain length, and held at a fixed potential[23] (Fig. 15.23). The equation resulting from treatment of the electrode kinetics of a packed bed (Fleischmann and Chu, 1974) in which it is assumed that deposition is controlled by transport processes is

$$\frac{c_L}{c_O} = e^{-L/\lambda} \tag{15.19}$$

where c_L is the concentration remaining after passage through the packed bed of length, L, and λ is a parameter that depends on the flow rate, the diffusion layer thickness at the given flow rate, and other factors.

If several metals are to be removed and separated, this can be done by redissolving the metals deposited on the bed (i.e., reversing the bed potential) and plating out each metal from the solution thus formed onto an electrode held by a potentiostat at specific potentials, each characteristic of a metal present in the solution; these can then be separately plated out and recovered. Thus, the characteristic reversible potential for each metal allows the removal of first those with the most positive potentials on the normal hydrogen scale and then, successively, those metals with increasingly less positive values for their reversible potentials. It is possible to get a 99% separation if the reversible potentials are as much as 0.3 V apart, but it may be necessary to make the plating-out solution alkaline so that the competing deposition of H_2 is suppressed. Packed-bed electrolysis has been applied not only to the removal of metal ions from dilute solutions but also to electro-organic synthesis, bromide recovery from brine, and the synthesis of hydrogen peroxide (Qi and Savinell, 1993).

Aquifers are underground bodies of water which, in the best instances, are as pure as rain water. Some of them are huge in extent (e.g., the Ogallala aquifer lies under part of Texas and several other states). They are a prime source of water for irrigation.

Unfortunately, some aquifers are being contaminated by seepage of toxic substances from agricultural products (and in some cases leakage of improperly stored industrial wastes). The removal of such contaminants by packed bed electrolysis applied to the exit pipe from the aquifer would allow continued use of these waters.

---

[23]The potential that is fixed is that of the metallic particles making up the bed. However, there is an IR drop through the solution which increases with the length of the bed, so that the potential available for deposition becomes significantly less as the length of the bed increases. This limits the effective value of L in Eq. (15.19).
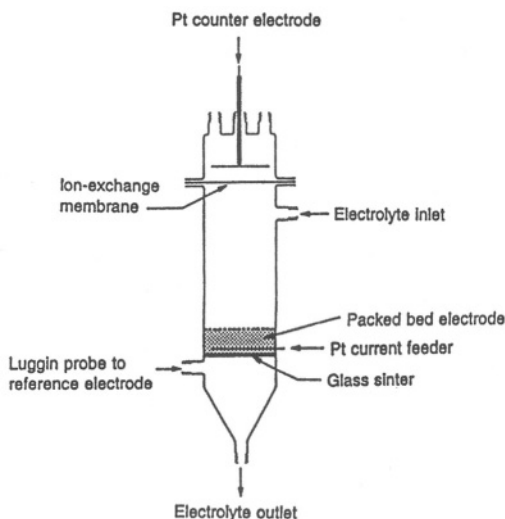
Fig.15.23. Packed-bed cell. (Reprinted from A. K. P. Chu, M. Fleischmann, and G. J. Hills, *J. Appl. Electrochem.* **4**: 324, copyright 1974.)

15.7.5. The Destruction of Nitrates

Very large quantities of radioactive waste materials are stored at various sites throughout the world.[24] The permanent disposal of these wastes involves dividing the material into two parts for treatment. The high-level radioactive species will be absorbed in porous solids and transported to depositories in salt mines or within mountains.[25]

There remain the so-called low-level wastes containing, among other toxic substances, dissolved Ru, Hg, and Cr, which, in volume, at the Hanford facility in Washington state alone, exceeded 100 million gallons in 1998. These materials are in the form of nitrates and present a hazard to the ecology and human health because if any disposal scheme allowed the nitrates to reach the ground water, they would contaminate the water. In surface water they promote the growth of unacceptable plants, such as algae.

The electrochemical reduction of the dissolved nitrate, eventually to NH_3 or perhaps even N_2, would remove this hazard (Hobbs and White, 1992). Of the metals

---

[24]In the United States these sites are in Hanford, Washington, and Savannah River, Georgia.

[25]It is claimed (Gleason, Fox, 1998) that electrochemically assisted nuclear reactions cause a 99% remediation of thorium to nonradioactive products. Should it be possible to apply such processes more widely, an improved solution to the waste problem may have been found.

present in the nitrate, those economically valuable would be recovered and the rest stored as oxides or hydroxides. On Pt and Ni electrodes, respectively, in solutions of NaOH containing $NaNO_3$, there is a cathodic peak at -0.6 NHE ($50\,mV\,s^{-1}$) on the Pt and at -0.7 on Ni. The steady-state (potentiostatted) curves for similar solutions are shown in Fig. 15.24.

At a Ni electrode, two different Tafel slopes were obtained as 30 mV/dec in the region from -0.7 to -0.77 V vs. NHE and 180 mV/dec in the region from -0.77 to -1.0 V vs. NHE (Fig. 15.24). Since NH_3 was produced as a product at the potential of -0.7 V vs. NHE, the two different Tafel slopes seem to signify two different rate-determining steps rather than two different reactions. In the region from -0.7 to -0.77 V vs. NHE, the transfer coefficient is 2. This coefficient can be represented as

$$\alpha = \gamma/v + m\beta$$

where γ is the number of electrons passing before the rate-determining step (rds), v is the stoichiometric number of the rds, m is the number of electrons involved in the rds, and β is the symmetry factor. A partial mechanism consistent with the observed slope and the stoichiometric number is given in Eq. (15.20). Further reduction products of NO are omitted because the intermediates after the rate-determining step do not affect the i–V curve.

$$NO_3^- + H_2O + 2e^- \rightleftharpoons NO_2^- + 2OH^-$$

$$NO_2^- + H_2O \rightarrow NO^+ + 2OH^- \text{ (rds)}$$

$$NO^+ + e^- \rightleftharpoons NO \qquad\qquad (15.20)$$
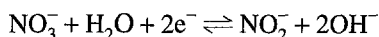
Such mechanism studies contribute to the aim of forming NH_3 or N_2 by reduction of NO_3^- because they provide the basis for a search for a good electrocatalyst that would accelerate the rds of the reduction down to the most desirable end product, molecular N_2. Ammonia might also be acceptable if it were filtered free of radioactive contaminants and therefore made commercially useful (Kim, 1997).

15.7.6. Electrochemical Treatment of Low-Level, Nuclear Wastes

The principal metals present in low-level nuclear waste nitrates are Hg^{2+}, Ru, in the form of a nitro complex, and CrO_4^{2-}. They are contained in a 1.3 M NaOH solution. The laboratory version of the basic device for their separation and removal is the packed-bed electrode shown in Fig. 15.25.

The concentrations of the metals are small and hence their electrochemical kinetics will be under transport control. The corresponding kinetics for the NO_3^- are activation control. A batch process would be used. The metals would be separated by controlling the potential in the packed bed, taking into account the change in potential
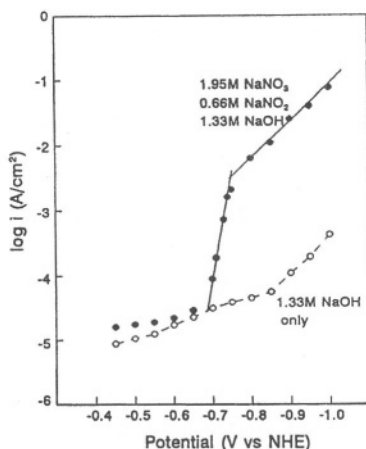
Fig. 15.24. Log i–potential plots for the reduction of a solution containing 1.95 M NaNO$_3$, 0.66 M NaNO$_3$, 1.33 M NaOH (solid circles) and of a solution containing 1.33 M NaOH only (open circles) at Ni. (Reprinted from J. O'M. Bockris and J. Kim, *J. Electrochem. Soc.* **143**: 3807, Fig. 14, 1996. Reproduced by permission of The Electrochemical Society, Inc.)

throughout the bed itself.[26] After the deposition of each metal, interruption of the flow process and current reversal would allow removal (and recovery for the Ru) of each metal.

The nitrate would be electrolyzed in a 24-hr cycle in parallel-plate cells with proton-exchange membranes. The optimal overall NO_3^- reduction reactions are

$$NO_3^- + H_2O + 2e^- \rightarrow NO_2^- + 2OH^-$$

$$NO_2^- + 5H_2O + 6e^- \rightarrow NH_3 + 7OH^-$$

$$2NO_2^- + 4H_2O + 6e^- \rightarrow N_2 + 8OH^-$$

---
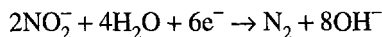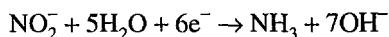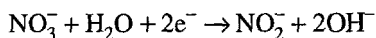
[26] The assumption that the potential is uniform in the metal of the bed but suffers an IR drop in the solution may be insufficient. Contact resistances between the metal particles set up a significant resistance and hence IR drop in the metal of the bed itself (Kim, 1997).

The experimentally determined dependence of the concentration of the various entities as a function of time is shown in Fig. 15.26. The catholyte and anolyte concentrate NaOH and HNO_3, respectively, using a membrane separator. A plan for the electrochemical treatment of low-level nuclear wastes is shown in Fig. 15.27. The considerable electricity costs of such processes could be compensated by the sale of HNO_3 and NaOH. The Ru might be commercially valuable for some purposes, but its use may be compromised by residual radioactivity.

15.7.7. Mediator-Aided Destruction of Organic Wastes (Particularly Toxic, Organic Waste)

Work in which numerous noxious organics were converted to CO_2 in electrochemical cells using mediators began to be published as early as 1975 (Clarke and Kuhn). Such processes work indirectly. An anode of, for example, Fe, Ni, or Ag produces the corresponding ions in the high valent state. In solution, these ions carry out oxidation processes with the wastes, and the resulting ions in the lower valent state
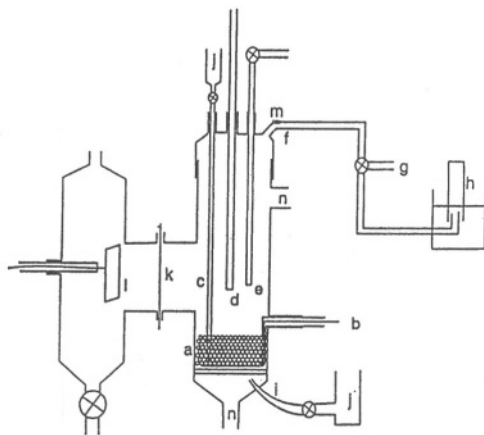
Fig. 15.25. Schematic for a laboratory-scale packed-bed electrode. a, Bed of particles; b, current collector; c, Luggin capillary; d, thermometer; e, purging gas in; f, gas out; g, bubbler; h, gas collector; i, Luggin capillary; j, reference electrode; k, Nafion film; l, counter-electrode; m, septum for gas analysis; and n, solution flow-in or flow-out. (Reprinted from J. O'M. Bockris and J. Kim, *J. Appl. Electrochem.* **27**: 625, copyright 1997.)
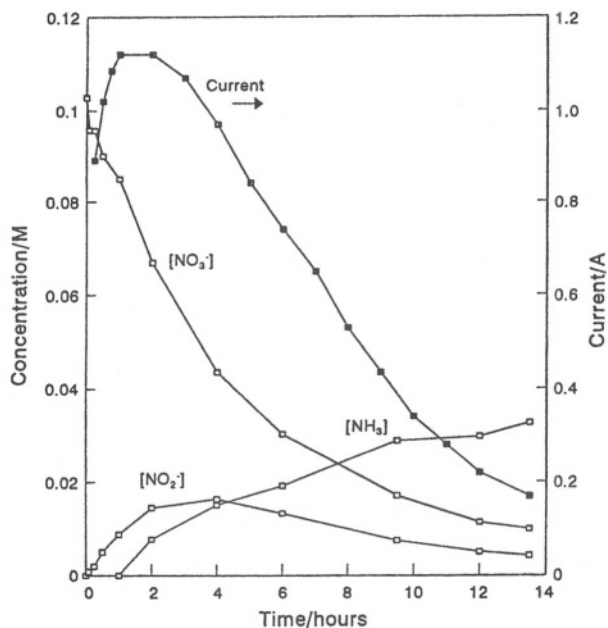
Fig. 15.26. Variation in the concentrations of nitrate, nitrite, annnonia, and the current as a function of time in the reduction of a solution containing 0.1 M $NaNO_3$ and 1.33 M NaOH at an Ni packed-bed electrode (cross-sectional area: 1.3×10^{-3} m^2; bed length, 0.02 m) at −0.7 V vs. NHE. (Reprinted from J. O'M. Bockris and J. Kim, *J. Appl. Electrochem.* **27**: 627, copyright 1997.)

are then reoxidized by the electrode and returned to the solution to accomplish further oxidation. Although this mechanism seems to be the right one in some cases, it will be in competition with direct oxidation.[27]

Surprising, even amazing, destructions can be achieved by this form of electroincineration. Coke, sewage sludge, and phenol have been treated under pressure at ~200 °C with Fe^{3+} as the oxidizing entity (Clarke 1990). Rubber gloves, gaskets, epoxy resins, kerosene, and oil wastes were oxidized away by using Ni and Ag ions (Steele, 1990).

---

[27]The mechanisms are easily distinguished. The homogeneous mechanisms lead to rates that are proportional to the concentrations of the mediator ions in solution and continue after the current is switched off (until the high valency form of the ions is consumed). The direct oxidation process stops when the current is turned off.

Fig. 15.27. A schematic of a possible plant for the electrochemical treatment of low-level nuclear waste. (Reprinted from J. O'M. Bockris and J. Kim, *J. Appl. Electrochem.* **27**: 626, copyright 1997.)

The advantage of electro over thermal incineration is that no noxious materials are added to the environment. It seems likely that rubbish disposal in general will become electrochemical. Low-temperature (~250 °C) molten salt processes may be necessary for tougher cases (Lin, 1997). Low current density Cl_2 evolution may be introduced as a biocide when necessary, particularly in treatment of aqueous electrolytic sewage involving urine (Kaba, 1990; Tennakoon, 1993).

Electroincineration could be used to deal not only with waste organics but also with the herbicides used in warfare (e.g., Agent Orange), explosives, and particularly

agents intended for use in biological warfare. Attempts to *burn* Agent Orange on a ship far out at sea still left 0.1% residues: 10 tons out of a 10,000 ton load!

15.7.8. Bactericidal Effects

Several electrochemical processes for the purification of wastes have been developed since 1975. The passage of liquids containing bacteria as well as larger protozoa between two electrodes was shown to destroy such life at 1 cycle s^{-1} and a current density of $10 \, \text{mA} \, \text{cm}^{-2}$. The processes occur in the solution phase, not at the electrodes (Stoner, 1975). Correspondingly, passage of water through a packed bed held at potentials that evolve H_2O_2 has been shown to sterilize the liquid (Gotto, 1991). Although electrochemical methods of sterilization show advantages over simply boiling the liquid, they may sometimes be in competition with UV irradiation.

As indicated earlier, metals in contact with sea water collect a slime consisting of the bodies of bacteria. If the pipe surfaces are pulsed in a potential region in which H_2O_2 is produced from dissolved O_2, the former compound acts bactericidally in the solution and the bacteria do not survive (Dhar, 1981).

15.7.9. The Special Problem of H_2S

15.7.9.1. *Introduction.* Bacterial processes convert organic sulfides and sulfates to H_2S in nature. Such processes result in H_2S being mixed with natural gas ("sour gas"), and if this is more than 10%, the gas is no longer considered to be commercially usable.[28]

15.7.9.2. *Electrochemical Decomposition of H_2S.* The reversible potential for the decomposition of water is 1.23 V, but that for the reaction

$$S + 2H^+ + 2e \rightarrow H_2S$$
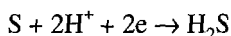
is 0.14 V at 25 °C. The thermodynamic reversible potential of a cell involving this reaction run in reverse at an anode with a corresponding hydrogen evolution at a cathode at pH 14 comes to 0.53 V at 25 °C. Thus, the production of H_2 from an alkaline solution of H_2S could proceed on the basis of a thermodynamic (minimum) energy expenditure that is 43% of that for the corresponding water decomposition of 1.23 V. The overpotential associated with the oxidation of sulfide to sulfur is less than that of oxygen evolution from water; and sulfur is of much greater commercial value than oxygen.

---

[28]Smaller quantities of H_2S can be removed by use of the Claus process:

$$3H_2S + \frac{3}{2}O_2 \rightarrow \frac{3}{x}S_x + 3H_2O$$
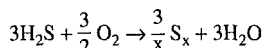
However, the useful H_2 is trapped inside the steam.

The cell potential for such a process is shown in Fig. 15.28 (Dandapani, 1987). In fact, at about 80 °C and pH 14, polysulfides are first produced, but pure sulfur precipitates on cooling. The process has been developed to an engineering stage (Petrov and Srinivasan, 1996). It is necessary to protect the cathode against catalytic inactivation by polysulfide (i.e., a membrane must be used). For continuous use, the pH must be controlled. The cooling and production of S may be best carried out outside the cell.

Thus, an electrochemical process for H_2 recovery from H_2S occurs at practical current densities of about half the energy needed for the recovery of H_2 from water. H_2S is plentiful the world over. Insofar as a market exists for the S, the income from the sale of by-product would more than compensate for the energy costs of the process, leading to negative hydrogen costs.
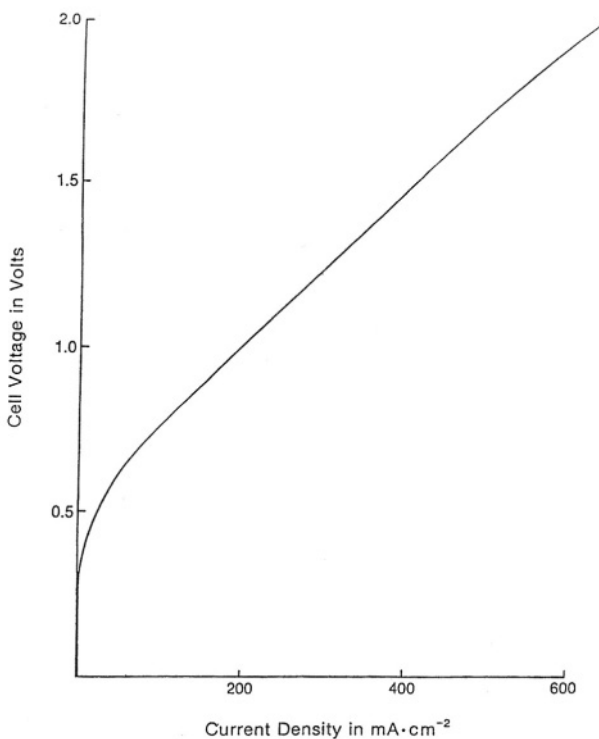
Fig. 15.28. Potential–current density relation due to the electrolysis of H_2S in an alkaline solution (70 °C) (Reprinted with permission of B. Dandapani.)

15.7.9.3. *Photoelectrochemical Decomposition of H_2S.*   The   photo-electrochemical decomposition of H_2S on CdSe with solar light alone has an efficiency of only 1.8% (Kainthla, 1987). However, photoelectrocatalytic processes and the use of two-electrode photocells would be expected to greatly increase the efficiency of conversion.[29]

15.7.10. Electrochemical Sewage Disposal

Sewage treatment plants end up with a sludge. What can be done with it? It may be spread over the land, whereupon rain will wash the toxic metals it contains (Hg, Cd, Pb) down to the water table. Alternatively, it can be burned, distributing the products of incomplete combustion far and wide onto other peoples' land and onto their heads.

Kaba and Hitchens (1989) found that electrolysis of a mixture of urine and feces produced CO_2, N_2, and H_2. Some HOCl is generated; this eliminates the pathogens and bleaches the contents. The anodic reactions at 90 °C consume the biomass; the cathode evolves hydrogen and can be assumed to deposit the small metal content. The residuum is sodium chloride from the urine. Most of the electrochemical studies that establish the basis of a practical process for electro-chemical sewage treatment have been carried out on packed-bed electrodes as shown in Fig. 15.29.

The anode material is important in the successful oxidation of sewage. If the flow rate through the packed bed is above a certain minimum, the anodic reactions are no longer transport controlled (because at higher flow rates the limiting current density is increased), and the electrochemical oxidation is subject to electrocatalysis. A suitable material base for the anode particles is reduced TiO_2 (Ebonex). A coating of SnO_2 greatly improves performance (Kötz, 1991). This can be further raised by doping the SnO_2 with Sb_2O_3.

A very small particle size for the fecal solids within the slurry is a critical condition upon which efficient consumption depends. The particle size can be reduced by the use of intense ultrasound to less than $25~\mu m$ in a device outside the cell. In a batch process, it has been shown (Tennakoon and Bhardwaj, 1996) that the total organic carbon can be reduced by >80% in 24 hr. The cost per person per day would be about 50 cents if electricity were available at 5 cents/kW hr^{-1}.

This direct approach to processing sewage (avoiding the preliminary sewage treatment plant and residual sewage sludge) eliminates the undesirable environmental impact of the combustion process, which is at present the most used process in dealing with sewage sludge. The direct electrochemical process must compete with the electrochemical mediator approach (Section 15.7.7).

---

[29]Namen (1997) has used colloidal suspensions of T_iO_2 loaded with Pt and RuO_2 to oxidize H_2S upon photoillumination. Such a method makes the separation of H_2 and S more difficult.
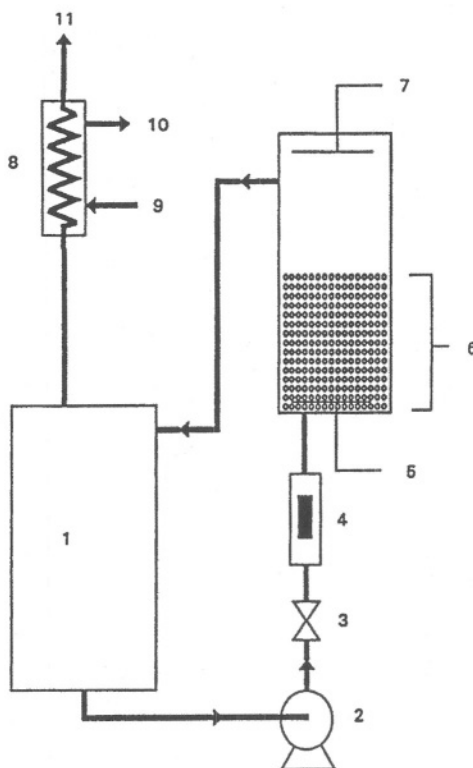
Fig. 15.29. Schematic diagram of a packed-bed cell and flow circuit. 1, Reservoir; 2, pump; 3, valve; 4, flowmeter; 5, anode current collector; 6, packed-bed anode; 7, cathode; 8, water condenser; 9, water inlet; 10, water outlet; and 11, outlet for gases. (Reprinted from C. K. Tennakoon, *J. Appl. Electrochem.* **26**: 27, copyright 1996.)

Some potential uses of the direct approach[30] are worth mentioning. The present process would be advantageous in recreational vehicles and boats, ships, and planes. Stand-alone sets for isolated communities could be powered by photovoltaics. In earth-based use, electricity costs could be reduced by the use of oxygen cathodes, thus also eliminating the evolution of hydrogen.

---

[30] The direct method for sewage disposal was originally developed for use on space vehicles where the hydrogen would be transduced in fuel cells to on-board electricity. The CO_2 evolved would be used for growing plants on board.

15.7.11. Electrochemical Decontamination of Soil

15.7.11.1. *Introduction.* One of the consequences of industrialization is that soils near factories tend to become contaminated. Such contamination occurs, for example, in Silicon Valley, California, where a number of plants and laboratories produce semiconductor materials. Waste containers (holding, e.g., compounds of Si, Ga, As, Zn, Cd, and Pb) sometimes overflow. Mismanagement (or benign neglect) allows the contaminating ions to reach the surrounding soil, where they gradually spread, like a stain on a fabric. Heavy fines are imposed by U.S. Environmental Protection Agency inspectors when such effects are detected.

The present means for decontaminating the soil involve removing it to a facility where it can be heated to a temperature that will decompose organic contaminants (250–400 °C) but also to the much higher temperatures that will melt or even vaporize some of the metal contaminants (the main body of soil is siliceous). The costs of such treatments are in excess of $1000/ton. The costs are high because the soil has to be dug up, transported to a treatment facility, and after incineration returned to the original site. An *in situ* method of decontamination is therefore highly desirable.

The first stages of an electrochemical remediation would involve surrounding the contaminated soil with a number of "stanchion"-type electrodes placed within the soil to a suitable depth, with a central counter-electrode, and applying an electric field between the two types of electrodes. Order-of-magnitude calculations show that currents would have to be passed for several months to remove 90% of the contaminants. Clearly, the soil must be kept wet.

15.7.11.2. *The Mechanism.*   A closer look reveals that the processes in such an operation involve electro-osmosis (Section 6.11.3). The contaminants in the soil are adsorbed on the surface of the soil particles, and there is an equilibrium between what is adsorbed and what is dissolved in the aqueous solution in the pores of the soil. One can present the electrochemical processes in a schematic form as shown in Fig. 15.30.

Around the cathode, OH^- ions will gather and tend to migrate away toward the anode. The anode will generate protons back into the solution. Some of these will simply migrate toward the cathode. However, some may displace the adsorbed M^+ ions (the contaminants), which can be seen in the diagram. M^+ ions dissolve to some extent and stretch out into a diffuse layer 10–1000 Å in thickness, depending on concentration.

On application of a field across "the cell," the anions (assumed in the diagram to form part of siliceous oxy ions constituting the soil) stay put, but the M^+ ions are pulled by the electric field toward the cathode (i.e., they are separated from the soil to which they formerly adhered). The degree of removal will be influenced by the solubility of the M^+ ions, and the type of complexation (if any) that may occur when they enter the solution. Remaining with the simplified picture of Fig. 15.30, one can see two ways
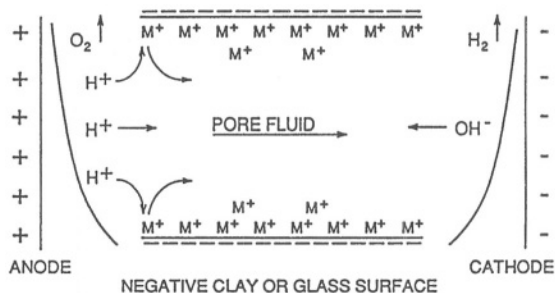
Fig. 15.30. Schematic of electrolysis, adsorption/desorption, and electro-osmotic flow. (Reprinted from R. J. Gale, "Soil Decontamination Using Electrokinetic Processing," in *Environmental Oriented Electrochemistry*, C. A. C. Sequeira, ed., Fig. 2, p. 623, copyright 1994. Reproduced with kind permission of Elsevier Science-NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.)

in which M^+ will move toward the cathode (or an anionic metal complex ion toward the anode).

1. For thinner pores, where the pore radius is small compared with $1/\kappa$, the Debye length of the diffuse layer, the dominant mechanism will be the electro-osmotic model of Fig. 15.30. The removal of the ions due to the transport of the diffuse-layer contents (cations predominating) toward the cathode will occur.

2. For thicker pores, in which the pore radius $R > 1/\kappa$, the electro-osmotic processes of (1) will play a part, but to this may be added ionic migration through the solution of ions outside the thickness of the diffuse layer.

Water will also flow. Thus, the cations and anions bring their hydration waters with them and since the hydration number (Vol. 1, Section 2.7) of a cation generally exceeds that of an anion, the net water transport will be toward the cathode. Apart from this effect as the origin of flow, there is a drag effect of the ion on its secondary solvation waters, the motion of which moves water outside the primary solvation sheath.

15.7.11.3. Experimental Work. The type of cell used in laboratory experimentations on these soils is shown in Fig. 15.31. Using apparatus of this kind, it has been possible to show a 30% removal of Cu^{2+} from sand in 30 days. Zn (in the form of zincate) has been removed from clay to a degree of 60% and Pb^{2+}, Cd^{2+}, and Cu^{3+} have been removed from soil to degrees varying from 60 to 90%.

A very important case is contamination of soils by radionuclides. Here, uranyl ions have been removed satisfactorily from some types of clay, but radium does not form an ion of sufficient solubility to allow electrochemical removal.
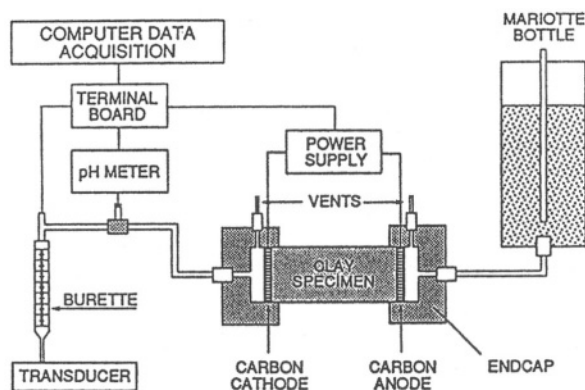
Fig. 15.31. One-dimensional laboratory test apparatus. (Reprinted from R. J. Gale, "Soil Decontamination Using Electrokinetic Processing," in *Environmental Oriented Electrochemistry*, C. A. C. Sequeira, ed., Fig. 4, p. 362, copyright 1994. Reproduced with kind permission of Elsevier Science-NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.)

15.7.11.4. *Summary on Soil Remediation.*[31]   Electrochemical   soil remediation is clearly an emerging technology and far from maturity. To the simplistic idea of removing contaminants from soil by electrolyzing the solution are added numerous effects that complicate the basic picture. What is the equilibrium constant of the ions concerned (some of which are adsorbed on the soil particles) with water? How much is electro-osmotic flow (Section 6.11.3) and how much is ionic migration in the solution? In what form are the ions present in this solution within the pores? How much will pH charges near the electrode affect wet particles contained among the processes, which may have to run for months? To all this may be added competing bacterial reactions.

In spite of these complexities, there *is* an electrochemistry of soil treatment and it has worked quite well in some cases. It exists not only in decontamination work but also in dewatering (electrodrainage). Thus, a mountainous mass of soil, say from mining excavation, may be in danger of starting to flow under gravity. In the past such flows have rolled down onto villages with fatal consequences. Clearly, the tendency to flow depends on the viscosity of the system and how this is affected by its water content. Insertion of a series of electrodes into the potentially mobile soil mass and a current flow could be used to remove water from the soil and reduce its flow properties.

---

[31] The material of this section owes much to the writings of Gale (1994).

15.8. RETROSPECT AND PROSPECT

The topics of this chapter have no lengthy past. In terms of decades, 1970 would be a reasonable beginning date on an imaginary plot of progress toward cleanup as a function of time. Since that year, it might be fairly said that much has happened, in spite of some lack of enthusiasm for the necessary major changes from powerful quarters associated with the oil-automotive imperium. In 1970 the Environmental Protection Agency was created; 1974 saw the formation of the International Society for Hydrogen Energy; in 1977 the National Renewable Energies Laboratory was established. However, the event that tops all these in its implications is the decision in 1997 by the Daimler-Benz Company in Germany to produce more than 100,000 fuel cell-driven passenger electric cars per year early in the new century. The on-board re-forming of methanol proposed by Daimler-Benz to obtain the hydrogen needed by the fuel cells is not the final solution (because the re-forming reaction will still produce some CO_2), but it is a major step in the right direction and is being followed in various forms by the principal automotive manufacturers.

Progress in the reclamation of some of the vast CO_2 content of the atmosphere has already had two peaks: the finding in 1992 of suitable catalysts for CO_2 reduction by H_2 to methanol and the publication of a detailed engineering analysis of an electrochemical extraction system for atmospheric CO_2. If these scientific advances resulted in commercial development, one could see a solution to automotive pollution, although extra H_2 is necessary to form CH_3OH, and this must come from electricity provided by renewable resources electrolyzing water. Thus, extraction of CO_2 from the atmosphere, conversion to methanol (using the electrolysis of water to obtain H_2) and re-forming to hydrogen could occur at gas stations (with no further need for a distribution system). The critical point here is that the carbon in the methanol should be taken back from the atmosphere because the re-forming of methanol to give hydrogen injects CO_2 into the atmosphere. In this way the main source of planetary warming would be permanently eliminated, and with it the need for fossil fuels.

In waste disposal, electrochemical (mediator) methods already have a foothold in European corporations. The concept of the anodically produced and reoxidized mediator has been proven and now needs scaleup and widespread introduction to municipalities throughout the world so that the incinerated products of present wastes no longer reach the atmosphere. Although the scope of the method already proven is remarkable, there remain problems of, e.g., the disposal of plastics and rubber tires. Direct methods (including the use of molten media) and those involving photoelectrochemical reactions are areas which await the funding of research to take the work from promising university laboratories to an engineering stage and perhaps to commercialization.

H_2S remediation in the oil industry, and in natural gas technology, seems to have found an economically superior process in the electrochemical version produced in the Center for Electrochemical Systems and Hydrogen Research at Texas A&M University. However, its use as a cleanup tool to replace the chemical Claus process

used at present may not be the most significant result of this development. The practical cell potentials at which H_2 can be produced from alkaline H_2S systems are about half those needed for the electrochemical decomposition of water. Insofar as there is a market for elemental sulfur, the returns from the sale of this product would more than compensate for the cost of producing hydrogen. Moreover, the availability of H_2S throughout the world is plentiful, so that its relatively cheap electrochemical splitting could increase the rate at which general conversion to a clean hydrogen-based energy system might proceed. A low-cost clean fuel (with no CO_2 injection into the atmosphere) would be hard to resist.

There are two electrochemical developments in sewage disposal. On the one hand, sewage sludge has been shown to be treatable by mediators (Section 15.7). On the other, the direct electrolysis of suspended microparticles originally designed for use in space vehicles has an advantage in that it coproduces H_2 and CO_2. Now that the catalytic means are available for forming methanol from H_2 and CO_2, direct automotive fuel production could be coupled to the electrochemical processing of sewage and indeed of all carbonaceous wastes treated by direct electrolysis and yielding CO_2.

An electrochemical process for treating low-level nuclear wastes is in the laboratory stage and not yet engineered. It must be understood less as a solution to the hazards of the low-level nuclear products contained in nitrate solution than as a solution for the remediation of the huge quantities of nitrate stored at nuclear waste depositories. The economic viability of the electrochemical process for conversion of nitrates to NH_3 and N_2 depends on the commercial development of its by-products, NaOH and HNO_3.

Finally, soil remediation processes are farther along than the remediation of low-level waste because some practical tests have been made in the field. In some systems, the results are those desired. The costs of the thermal process for removing metal from soils are so high that funding of the development of the electrochemical approach would seem to offer good returns.

One large area has been omitted here, that of detecting and monitoring pollution by electrochemical methods. Much of this material is described in texts on electroanalytical chemistry.

15.9. A PARTING WORD

Something must also be said on the socioeconomic side and how it affects the rate of progress toward clean, electrochemical technologies. Our economic system, which now dominates the world, runs on short-term profits. The perception of profit and return on investment is geared to a time scale (2–3 years) that is far less than the decades needed for innovative research and development. While this short-term view ensures quick profits and makes executives look good to stockholders, it means that there is no planning for the long term. And it is the long-term view that is required for a stable, viable future.

The path of an idea for a new, improved technology to commercial availability is a long one, often requiring six to eight decades and substantial funds. This is rarely a quick process. European and Asian governments have long recognized the important role they can play in supporting research in critical areas. This is not true in the United States, where funding decisions are subject to partisan politics and often are made by persons who have a financial interest in maintaining the status quo.

Many of the electrochemical ideas and methods described in this chapter are in various stages of development. Some are assured of commercialization (e.g., electrochemical transportation); some are in industrial development but are not yet well known (e.g., electrochemical destruction of wastes); and some are still in university research laboratories, waiting for funding (e.g., remediation of nitrates from nuclear wastes). And while they wait for funding, environmental degradation from all sources continues, along with the accumulation of greenhouse gases.

It is not satisfactory to rely on the short-term profit motives of corporations to provide the funding needed, or on the representatives of these corporations who sit on funding committees.[32] The revolutionary technologies that are needed to ensure a sustainable economy and an undegraded quality of life in this country in the twenty-first century cannot be created without a serious, nonpartisan commitment from a government that takes a long-term view of the well-being of the population and not the short-sighted one of corporate executives.

Further Reading

1. J. Tyndall, *Phil. Mag.* **22**: 169, 273 (1861). First publication on the greenhouse effect.
2. F. Fischer and O. Prziza, *Ber. Deutsch Chem. Ges.* **47**: 256 (1914). First electrochemical reduction of CO_2.
3. G. Plass, *Quart. Roy. Meteorolog. Soc.* **82**: 310 (1956). First calculation on the greenhouse effect.
4. Rachel Carson, *Silent Spring,* Houghton-Mifflin, Boston (1962). Some see this book as triggering environmental consciousness.
5. E. Justi, *Conduction Mechanism and Energy Conversion in Solids* , Udo Pfriemer Verlag Gottingen (1965). See particularly, Fig. 23, a schematic of a solar-hydrogen scheme.
6. J. O'M. Bockris, *Environment* **13**: 31 (1971). The first published paper proposing a general use of hydrogen as an energy medium.
7. J. O'M.Bockris, ed., *The Electrochemistry of Cleaner Environments* , Plenum, New York (1972). The first collection of papers dealing with a specific environmental problem in an electrochemical way.
8. K. A. Ehricke, *The Power Relay Satellite,* North American Aerospace Group, Rockwell International (December 1973). The concept that solar energy should be collected in highly

---

[32] This is so because corporations seldom support researchers on processes unlikely to yield a profit in less than 3 years.

insolated areas of the world and beamed in microwave radiation to a satellite, which would in turn beam it to distant continents.

9.   M. Fleischmann and A. K. O. Chu, *J. Appl. Electrochem.* **4**: 323 (1974). Theory of electrochemical extraction in a packed bed.

10. J. O'M. Bockris, *Energy: The Solar-Hydrogen Alternative,* Australia and New Zealand Book Company, Sydney (1975). The first book in which the use of solar energy to provide societal needs is presented at a technical level.

11. R. L. Clarke, A. T. Kühn, and E. Okoh, *Electrochemistry in Britain* (1975). Destruction of wastes via the mediator approach.

12. G. Stoner, U.S. Patent, 3,725,326, 1975. Electrochemical sterilization.

13. J. O'M. Bockris and K. Uosaki, *J. Electrochem. Soc.* **124**: 98 (1977). Stability of the electrodes in photoelectrolysis.

14. A. Szent-Gyorgyi, in *Submolecular Biology and Cancer,* p. 1, Excerpta Medica, New York (1978). Ideas on photons and solar energy storage in biomass.

15. V. Guruswamy and J. O'M. Bockris, in *Solar Energy Materials,* Vol. 1, p. 441, 1979. Hydrogen and electricity from water and light.

16. B. G. Pound, D. J. M. Bevan, and J. O'M. Bockris, *Int. J. Hydrogen Energy* **6**: 473 (1980). The electrolysis of steam to 1650 °C to form hydrogen.

17. V. Guruswamy, O. J. Murphy, V. Young, G. Hildreth, and J. O'M. Bockris, in *Solar Energy Materials,* Vol. 6, p. 43, 1981. Photon electrochemical production of Cl_2 from sea water.

18. H. P. Dhar and J. O'M. Bockris, *J. Electrochem. Soc.* **128**: 229 (1981). Elimination of bacterial coatings by electrochemical reduction of O_2 to H_2O_2.

19. O. Murphy and F. Gutmann, "The Electrochemical Splitting of Water," in *Modern Aspects of Electrochemistry,* R. White, J. O'M. Bockris, and B. E. Conway, eds., Vol. 15, p. 1, Plenum, New York (1983). Review.

20. J. Ghorogchian and J. O'M. Bockris, *Int. J. Hydrogen Energy* **10**: 101 (1985). Homopolar generator in the electrolysis of water.

21. Y. Hori, K. Kikucki, A. Murah, and S. Suzuki, *Chem. Lett.* **34:** 897 (1986). Reduction of CO_2 to MeOH on Cu electrodes at 0 °C.

22. R. C. Kainthla and J. O'M. Bockris, *Int. J. Hydrogen Energy* **12**: 23 (1987). Photoelectro-chemical decomposition of H_2S.

23. K. Chandresekaran and J. O'M. Bockris, *Surface Sci.* **185**: 495 (1987). FTIR evidence of CO_2^- ion as an intermediate in the reduction of CO_2.

24. I. Taniguchi, in *Modern Aspects of Electrochemistry,* J. O'M. Bockris, R. E. White, and B. E. Conway, eds., Vol. 20, p. 127, Plenum, New York (1989). Photoelectrocatalysis in the reduction of CO_2.

25. J. O'M. Bockris and J. Wass, *J. Electrochem. Soc.* **136**: 2523 (1989). Photoelectrocatalytic mechanism in CO_2 reduction using macrocycles.

26. M. Oppenheimer and B. Boyle, "Dead Heat," *New Republic,* New York (1990). A description of the consequences of the present trend in global warming.

27. D. F. Steele, *Platinum Met. Rev.* **34:** 10 (1990). Treatment of mixed wastes by Ag^+ and the mediator technique.

28. L. Kaba, G. D. Hitchens, and J. O'M. Bockris, *J. Electrochem. Soc.* **137**: 1341 (1990). Electrochemical destruction of sewage.

29. C. L. K. Tennakoon, R. C. Bhardwaj, and J. O'M. Bockris, *J. Appl. Electrochem.* **26**: 15 (1990). Packed bed use in electrochemical destruction of sewage.

30. M. Halleman, in *Proc. Int. Symp. on Chemical and Electrochemical Fixing of Carbon Dioxide,* paper A19, Chemical Society of Japan (1991). Reduction of CO_2.

31. M. Enyo, T. Atoguchi, and Akiko Aromata, *Proc. Int. Symp. on Chemical and Electro - chemical Fixing of Carbon Dioxide,* p. 333, Chemical Society of Japan (1991). Macrocycles in CO_2 reduction.

32. G. D. Hitchens, O. J. Murphy, L. Kaba and C. E. Verotsko, *20th Int. Conference, Environmental Systems* (1991). Electrochemical purification of waste water.

33. R. Kotz, S. Stucki, and B. Carcer, *J. Appl. Electrochem.* **21**: 14 (1991). Doped tin oxide anodes and their use in organic oxidation.

34. D. H. Meadows, D. L. Meadows, and J. Randers, *Beyond the Limits,* Chelsea Green, Post Mills, VT (1992). A 1992 update on the concept that there are limits in world resources and these will be reached in the twenty-first century.

35. K. Uosaki and S. Nakabayashi, *Chem. Lett.* **40**: 1474 (1992). Upgrade of chemicals by photoelectrochemical reactions with CO_2.

36. D. T. Hobbs, Technical Report on the Electrochemical Treatment of Alkaline Nuclear Wastes. DOE Report WSRC-TR 94-0287 (1994). Review.

37. R. Gale, in *Environmentally Oriented Electrochemistry,* C. A. C. Sequeira, eds., Elsevier, Amsterdam (1994). Electrochemical destruction of hazardous wastes.

38. S. Stucki, A. Schuler, and M. Constantinescu, *Int. J. Hydrogen Energy* **20**: 653 (1995). Extraction of CO_2 from the air by an electrochemical method.

39. Y. B. Acar, R. J. Gale, A. N. Alshawabkeh, R. E. Marks, S. Puppala, M. Bricka, and C. Parker, *J. Hazardous Mat.* **40**: 117 (1995). Basics and technological status of electrochemical soil remediation.

40. K. Petrov and S. Srinivasan, *Int. J. Hydrogen Energy* **21**:163 (1996). Chemical engineering of the electrolysis of H_2S to hydrogen and sulfur.

41. Y. B. Acar, F. Ozsu, A. N. Alshawabkeh, M. Rabbi, and R. J. Gale, *Chem. Tech.,* April 1996, p. 40. Biodegradation under electric fields.

42. Y. B. Acar and A. N. Alshawabkeh, *J. Geotech. Eng.* March, 1996, p. 173 and p. 186. Pilot-scale tests of lead removal from kaolinite.

43. J. O'M. Bockris and J. Kim, *J. Appl. Electrochem.* **27**: 623 (1997). Electrochemical treatment of low-level nuclear wastes.

44. J. O'M. Bockris and J. Kim, *J. Appl. Electrochem.* **27**: 890 (1997). Effect of significant contact resistance between particles in a packed bed on the distribution of current.

45. A. Namen, *Int. J. Hydrogen Energy* **22**: 783 (1997). Compares straight photoelectrolysis with photoirradiation of colloids.

46. D. Nagel, *J. Radiation Phys. Chem.* 1998. A review of low-energy nuclear reactions by a division chief of the U.S. Naval Research Laboratory.

47. F. DiMascio, J. Wood, and K. Fenton, *Interface* **7**: 27 (1998).

48. C. Platt, *Wired,* Nov. 1998, p. 172. Cold fusion is real.

EXERCISES

1. Let us consider a water electrolyzer characterized by the following: Anode: Ni, surface area $S = 10 \text{ cm}^2$; cathode: Ni, surface area $S = 10 \text{ cm}^2$; electrolyte: 1000 ml 40% NaOH, pH 14, T = 20 °C, distance between electrodes: $L = 10 \text{ cm}$; electrolyte conductivity $\sigma = 0.31 \ \Omega^{-1} \text{ cm}^{-1}$; diaphragm between the anodic and cathodic compartment: thickness $= 0.1 \text{ cm}$, resistivity $\rho = 10^3 \ \Omega \text{ cm}$; dimensions: $10 \times 10 \ (\text{cm}^{-2})$. (a) Write down the reactions occurring at the two electrodes; (b) In what compartment is water consumed and must be replaced? (c) Why is it necessary to add NaOH?

    The reaction must be driven to go against its spontaneous tendency. The standard enthalpy of water in the liquid state is $-56,690 \text{ cal mol}^{-1}$, and the equilibrium potentials of the oxygen and hydrogen electrode vs. NHE are $E^{\circ}_{H_2/H_2O} = -0.059 \text{ pH}$ and $E^{\circ}_{OH/O_2} = 1.23 - 0.059 \text{ pH}$, respectively. (d) Calculate the theoretical potential difference E_{th} between the two electrodes and the potential difference under open-circuit conditions ($i = 0$), to produce water decomposition. (e) Calculate the potential difference, t, to be applied between the two electrodes for a current density $i_{an} = -i_{cath} = 0.1 \text{ A cm}^{-2}$, taking into account that the oxygen overvoltage at Ni is $\eta_{an} = 0.26 \text{ V}$ and the hydrogen overvoltage at Ni is $\eta_{cath} = -0.11 \text{ V}$. (f) Calculate the quantity of molecular hydrogen and oxygen (in grams) produced after 24 hr of electrolysis and the quantity of water to be added. After every symbol and every numerical quantity write the dimension in parentheses. (Plonski)

2. Ten cubic meters of a nonbiodegradable industrial waste water with a chemical oxygen demand (COD) of 6 kg/m^3 has been pretreated by anodic oxidation. After 5 hr of electrolysis at 300 kA the total organic carbon (TOC) is 1 kg/m^3. Analysis of the waste water after electrochemical treatment shows that the final oxidation product was oxalic acid. Calculate: (a) The concentration of oxalic acid and the COD after the electrochemical treatment. (b) The average current efficiency and the anode surface used if the applied current density was 1 kA/m^2. (Cominellis)

3. A chemical plant produces 5 m^3 of a waste water with the following composition: phenol, 3 kg/m^3; acetone, 4 kg/m^3; Na_2SO_4, 25 kg/m^3, and H_2SO_4, 5 kg/m^3. (a) Calculate the TOC, COD, and the average oxidation state $(\bar{O}_x)$ of the organic carbon in this waste water before treatment. (b) Calculate the anode surface area necessary to eliminate 60% of the COD by anodic oxidation, using a current density of 2 kA/m^2 and considering an average current efficiency of 50%. (Cominellis)

4. The anodic dissolution of Fe or Al anodes is used for the electrocoagulation of organic and inorganic pollutants in waste water. If a total charge of 30,000 C is passed through an Al anode immersed in an electrochemical reactor containing 100 liters of waste water (assume a 50% current efficiency and plenty of hydroxide ions in it), would enough aluminum hydroxide be produced to satisfy a requirement of 30 mg of this hydroxide per liter of solution to clean it by this process? (Ibanez)

5. There have been severe criticisms about the extended use of chlorine gas in industry, owing to concern primarily derived from its ability to form toxic chlorinated organic compounds. In order to avoid its co-production during the electrolytic production of sodium hydroxide, a process has been developed in which a sodium carbonate (soda ash) solution is used as the anolyte in an electrochemical reactor divided by an ion-exchange membrane. Hydrogen gas is produced at the cathode and sent to a gas diffusion anode. Assuming no by-products in the liquid phase and only one by-product in the gas phase:

    (a) Draw a diagram of the process, (b) What type of membrane (i.e., cathodic or anodic) is required? (c) Write the balanced individual equations for the reactions that occur at each electrode and in the solutions of each compartment, (d) Write the global reaction. (Ibanez)

6. Potassium stannate, $K_2Sn(OH)_6$, is produced in the anodic compartment of an electrochemical cell at room temperature and without requiring high currents. This energy saving is also an environmental advantage. The tin anode is separated from the (inert) cathode by an-ion-exchange membrane. The applied potential is capable of only partially oxidizing Sn. The anolyte consists of an air-sparged KOH solution. The catholyte also consists of a KOH solution. Both Sn(II) and Sn(IV) form in the cell tetra- and hexa- hydro complexes, respectively, which in principle could be reduced at the cathode.

    Without further chemical or electrochemical information: (a) Sketch the cell and the processes occurring in it. (b) What is the purpose of aerating the anolyte? (c) What type of membrane (cationic or anionic) is required? (d) Write the balanced equations that describe the process at: the anode, the anolyte, and the cathode. (e) Write the balanced global equation. (Ibanez)

7. Formaldehyde is a toxic organic substance. It has been found that it can be degraded to harmless products with electrogenerated hydrogen peroxide either in acidic (with the aid of a catalyst, not shown below) or basic solutions according to the following equations, respectively:

$$O_2 + 2H^+ + 2e^- = H_2O_2$$

$$HCHO + H_2O_2 = HCOOH + H_2O$$

$$O_2 + H_2O + 2e^- = HO_2^- + OH^-$$

$$2HCHO + HO_2^- + OH^- = 2HCOO^- + H_2O + H_2$$
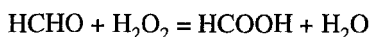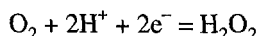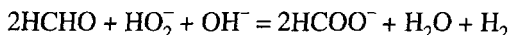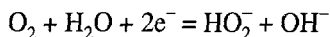
Based only on stoichiometry and assuming no side reactions, in which case will the current efficiency (defined as the charge employed for a given process divided by the total charge passed through the system) for formaldehyde destruction be higher? (Ibanez)

8. Tap water and swimming pool water have been disinfected by the addition of an aqueous solution of ClO_2, which can be produced as shown schematically in Fig. E15.1.
   Here R_1 and R_2 are chemical reactors (a) Identify the chemical nature of A, B, C, and D. (b) Write down the balanced reactions occurring at the anode, at the cathode, in reactor 1 (this is a disproportionation reaction), and in reactor 2 (the products are ClO_2 + HCl + NaCl).
   Note that H_2O does not react with $A_{(g)}$, but it does react in reactor 2. (Ibanez)

9. Cobalt recovery from acidic metal ion solutions is impeded by the evolution of hydrogen. However, if the pH of the solution is gradually increased, it is found that at pH > 4 this deposition is possible.
   (a) Explain this fact with the aid of the appropriate equations and sketch the corresponding E–pH diagram. (b) Give a reasonable estimate of $E^{\circ}_{Co^{2+}}$, in volts (vs. SHE) assuming standard conditions.

10. Mercury has been removed from contaminated brine solutions by means of a reticulated vitreous carbon porous electrode. A reduction factor as large as 5000 was obtained in a single pass. It was shown that the concentration ratio can be represented by the following equation:
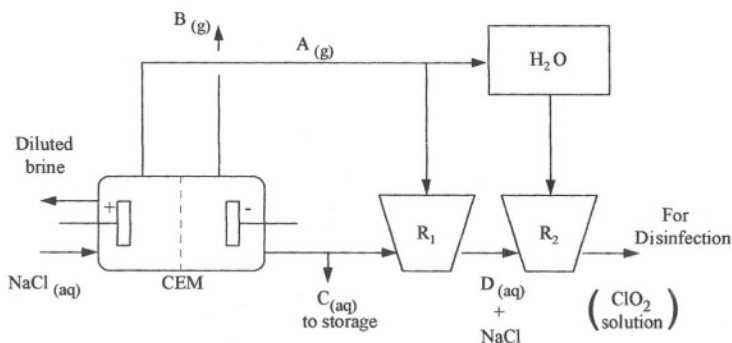
Fig. E15.1.

$$\frac{C_x}{C_0} = \exp\left(-\frac{k_m \times A_e}{v} x\right)$$

where C_0 and C_x are the initial mercury ion concentration and the concentration at an electrode length of x cm; k_m is the local mass transfer coefficient (assumed to be constant and equal to 2.1×10^{-4} cm s^{-1}); A_e is the electrode area per unit of electrode volume (66 cm^2/cm^3) and v is the fluid velocity (2.55×10^{-2} cm s^{-1}). With these data: Find the length (x) for which a 99.9% removal of the initial mercury is attained in a single pass. (Ibanez)

11. An industrial site polluted with lead to an average concentration of 3000 ppm is to be remediated using electrokinetic soil processing. The site, 60 feet long $\times$ 30 feet wide $\times$ 12 feet deep, is fitted with two rows of electrode wells along its length, 30 feet apart (6 anode wells, 6 cathode wells, 10 feet apart, 5 feet offset). Assume that a uniform, parallel electric field is created and the constant, cross-sectional current density is 200 μA cm^{-2}. (a) What is the total current required? (b) If the soil conductivity is 450 μS cm^{-1} at the start of processing and linearly rises to 900 μS cm^{-1} over 9 months, what is the average power required (kW hr/yd^3)? (c) The extraction efficiency was 86% and the soil wet density was 1.88 g liter^{-1}. (d) How much lead (kg) was recovered? (Gale)

12. A coastal beach in California is polluted with heavy metals. Since it is a protected wildlife habitat, a minimally intrusive electrochemical method is selected for cleanup. Assume that a constant current density of 125 μA cm^{-2} in a 40 $\times$ 6-foot cross section is used in the contaminant pit, which is 40 $\times$ 20 $\times$ 6 feet deep. (a) What is the total current and voltage required if the pore fluid conductivity is 21.9 mS cm^{-1} (approx. equivalent to 0.2 M KC1)? (b) If the soil is saturated and approx. 50% pore fluid and 50% solids by volume, how long would it take to pass a charge equivalent to the ionic content of the pore fluid? (c) How much acid should be added to depolarize the cathode in this time in order to ensure reaction (A) below, instead of water electrolysis, reaction (B)?

$$2H^+ + 2e^- \rightleftharpoons H_2 \tag{A}$$

$$2H_2O + 2e^- \rightleftharpoons H_2 + 2OH^- \tag{B}$$
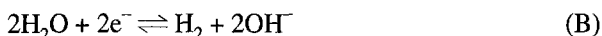
(d) What product might you expect to find at the inert anode?

PROBLEMS

1. The "greenhouse effect" (planetary warming) has been the focus of much interest by environmentally active scientists for decades, (a) What are the most probable temperature changes at the equator and at the poles for 2050? (b) What

range of values is given for the expected sea level rise at this time?

One solution to the problem of planetary warming (caused by the use of fossil fuels) involves the collection of solar light, its conversion to electricity, and the electrolysis of water to give hydrogen, the latter to be piped to cities as natural gas is now piped, for example, from Texas to New York.

(c) Calculate the area of the average solar farm if solar collection occurs in 20 states. The solar light is to be converted to electricity at a 20% efficiency. As a simplifying assumption, assume the availability of sun 12 hr per day for half the year. (Assume a levelized use of all types of energy as being equivalent in the United States to 10 kW/person. The solar collectors are assumed to be oriented orthogonally to the sun and to receive 1 kW m^{-2}.) (Bockris)

2. Although water electrolysis is the method of obtaining hydrogen free from CO_2, it has been claimed that the electrolysis of H_2S, in alkaline solution would give hydrogen at a significantly lower cost. Investigate this. (a) Make a schematic of the relevant electrochemical cell. (b) Write the equation for the reactions at each electrode (the anodic reaction would convert HS^- to S). (c) What would be the reversible cell potential at pH 14 and T = 25 °C? (d) Compare this value with that for the corresponding electrolysis of an alkaline solution to give H_2 and O_2. (e) Would the solution tend to change pH as the electrolysis proceeded? (f) Would a membrane be necessary and if so, what kind? (Handbooks will have to be consulted to obtain the data needed to solve this problem.) (Bockris)

3. If a full hydrogen economy were to develop in the twenty-first century, very large amounts of water would be necessary and the question arises as to whether highly conducting sea water would be usable. (a) Consider whether the danger of Cl_2 evolution could be avoided by electrolyzing brine (concentrated sea water) while keeping the potential of the anode restricted to that for the evolution of Br_2 (a liquid at 25 °C). (b) Find the reversible potential for the evolution of Br_2 (see the *Handbook of Chemistry and Physics*). (c) If brine concentrates sea water 5 times, what would be the Br^- concentration? (d) If a rotating anode were used with an rpm of 5000, what would be the limiting current for Br_2 evolution (use the Levich equation for δ_ω)? (e) If the cathodic current density is 100 mA cm^{-2}, calculate the size of the anode (rotating cylinder, 2500 rpm) so that Br_2 can be evolved below its limiting current density. (f) Calculate the potential at which the potentiostat would have to control the anode potential so as to avoid Cl_2 evolution. (g) What could be done with the excess liquid Br_2 so as to avoid any environmental hazard? (Bockris)

4. Aquifers are underground lakes of fresh water, some of them the size of smaller U.S. states. Some aquifers are showing signs of pollution. It is proposed that the flowing water should be purified electrochemically at exits to the aquifer. In this book, descriptions of packed beds have been given, as well as other means for electrochemically purifying polluted water. As a simplification assume that the

most important pollutant in the aquifer water is lead at a concentration of 1 ppm.

(a) Convert this concentration to moles per cubic centimeter. Consider electroreduction of this impurity at a flat-plate electrode in a stationary electrolyte. The basic equation (assuming diffusion control) is

$$\frac{d}{dt} Vc = \frac{iA}{nF}$$

where V is the volume of the liquid to be purified, i is the limiting current density of the electrolytic oxidation in amperes cm^{-2}, n is the number of electrons needed in one act of the overall reaction, and c is the concentration in mol cc^{-1}.

(b) Using the relation

$$i_L = \frac{DnFc}{\delta}$$

show that the concentration after t seconds is related to the original concentration by

$$\frac{c_t}{c_0} = e^{-D_{At}/\delta v}$$

Consider an extraction device cell of 100 parallel plates, (c) What would be the area of each plate needed to purify batches of a cubic meter of water from 0.1 to 0.001 ppm in 100 s? (Bockris)

5. In the normal water electrolysis cell, the reversible potential at 25 °C is 1.23 V, but the potential for practical electrolysis at >0.1 A cm^{-2} is >1.7 V. Thus the overpotential (~0.5 V) adds about 30% to the cost of H_2.

    Aqueous solutions all evolve H_2 when the cathode potential is made sufficiently negative. However, it may be possible to have an aqueous solution that contains inexpensively dissolved substances (e.g., SO_2) that become oxidized at potentials much less than that of water itself. Then, looked at thermodynamically, the reversible potential of the reaction in the cell would be less than that of water. In addition, the i_0 value for oxygen evolution ($i_0 \sim 10^{-10}$ A cm^{-2} at 25 °C) is particularly low and the anode overpotential particularly high. Substitution of, e.g., SO_2 oxidation could be achieved at a lesser overpotential than with O_2 evolution.

    (a) Examine SO_2 and NO as possible candidates, the dissolution of which in water might give hydrogen more economically than straightforward water electrolysis. Use handbook data on gas solubilities to obtain the limiting currents for these materials.

    (b) Write down the anode reactions that would be relevant to a practical realization of this concept. (c) Find out using a handbook the reversible electrode

potentials for the most likely oxidation in acid solutions. (d) Assuming pH = 0, calculate the reversible potentials of cells that evolve H_2 and use SO_2, and NO as the basis of the anodic reactions.

  (e) Finally, calculate the savings in cost by using these waste products (which are assumed to be supplied free of charge) compared with that of water electrolysis to H_2 and O_2: without accounting for sale of H_2SO_4 or HNO_3; accounting for such sales (count the anode overvoltage as half that for O_2). (Bockris)

6. In aqueous media, phenol can be oxidized on a fluidized bed (composed of Pb particles) that acts as the anode in an electrochemical reactor as follows:
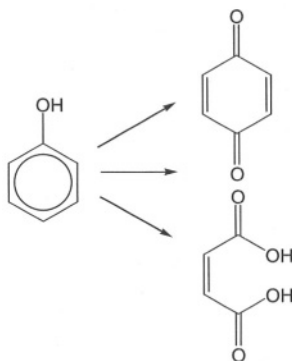
Assuming that these are the only products (plus water), and that the initial number of moles of phenol is n_{Ph_0}, write the mass balance that describes at any time the change in number of moles of phenol, $(n_{Ph_0} - n_{Ph})$ as a function of the number of moles of benzoquinone (n_{BQ}), maleic acid (n_{MA}) and carbon dioxide (n_{CO_2}) produced up to that moment as follows:

$$(n_{Ph_0} - n_{Ph}) = f(n_{BQ}, n_{MA}, n_{CO_2}). \text{ (Ibanez)}$$

7. An electrolytic process reduces the concentration of corrosive nitrates from a nuclear waste site more than 99%. The reduction of nitrate ions yields nitrite ions, which then are further reduced to produce the gaseous nitrogen-containing species: N_2O, N_2, or NH_3, depending upon process conditions. The corresponding equations are:

$$aNO_3^- + aH_2O + (2a)e^- = aNO_2^- + (2a)OH^- \tag{1}$$

$$aNO_2^- + bH_2O + ce^- = A + dOH^- \tag{2, 3, 4}$$
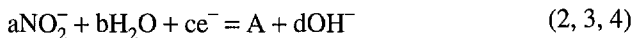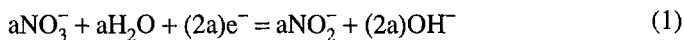
where a, b, c, and d are the stoichiometric coefficients necessary for the production of 1 mol of A (where A = N_2O in reaction 2, A = N_2 in reaction 3,

and A = NH_3 reaction 4).

The dominant reaction at the anode is known to be

$$4OH^- = O_2 + 2H_2O + 4e^-$$ (5)

(a) Find the values of a, b, c, and d for reactions 2, 3, and 4. (b) Write (and balance) the following net reaction for the reduction of nitrite ions for each of the three cases described above:

$$aNO_2^- + wH_2O = xA + yOH^- + zO_2$$ (6, 7, 8)
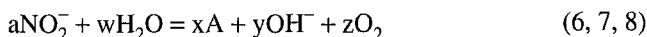
(here A = N_2O in reaction 6, A = N_2 in reaction 7, and A = NH_3 in reaction 8).

(c) Find the ratio of equivalents of hydroxide produced to equivalents of nitrate reduced (i.e., the ratio y/a) for the three cases described above (where A = N_2O, N_2, or NH_3). (Ibanez)

8. The general scheme for the electrocatalytic oxidation of an organic molecule ($C_nH_xO_y$) by Ag(II) ions in the form of $AgNO_3^+$ is as follows.

First, Ag(I) is oxidized to Ag(II) at the anode of an electrochemical cell. Then, Ag(II) oxidizes the organic to CO_2:

$$aAgNO_3^+ + C_nH_xO_y + bH_2O = aAg^+ + nCO_2 + aHNO_3$$
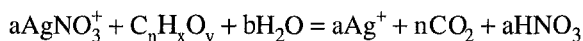
and the cycle can be repeated again. It is important to note that the more an organic species is oxidized, the fewer electrons need to be removed to achieve its complete conversion to CO_2. For example, for the oxidation of the two carbons of acetic acid to CO_2 (where the carbon atom is in an oxidation state of +4), eight electrons need to be removed, whereas the same procedure for ethanol requires removal of $12\ e^-$. With this information: (a) Write the above equation for the following substances and balance the resulting equations: ethylene glycol, acetone, and benzene. (b) Give the formula for at least two substances of the type $C_nH_xO_y$ for which $n = b = 3$, and write the corresponding balanced equations to describe their electrocatalytic oxidation as described above. (Ibanez)

9. Two cubic meters of a nonbiodegradable organic industrial waste water with a COD of 10 kg/m^3 and a TOC of 5 kg/m^3 are to be pretreated (before the biological treatment) by anodic oxidation under galvanostatic conditions ($i = 1$ kA/m^2) using a filter press electrochemical reactor with a 20 m^2 anode surface area. Consider that:

After elimination of 60% of the COD the waste water becomes biodegradable.

The average current efficiency for eliminating 60% of the initial COD is 30%.

The average cell potential during the electrochemical treatment is 4.5 V.

Calculate the treatment time and the specific energy consumption (kW h/kg COD eliminated). (Cominellis)

10. A continuous electrochemical plant is to be designed for the treatment of $2 \ m^3/hr$ of an organic industrial waste water with a COD of $2.5 \ kg/m^3$. Consider that:
    The treatment has been carried out at constant current density ($i = 1 \ kA/m^2$).
    The average current efficiency for the elimination of 80% of the COD is 35%.
    The average cell potential during the electrochemical treatment is 5 V.
    Calculate the required anode surface area and the specific energy consumption (kW hr/kg COD eliminated). (Cominellis)

11. It is necessary to evaluate the injection homogeneity of a carbon source (cometabolite) for microbial degradation enhancement. It is injected electro-osmotically at 4 cm/day at 100 ppm with a homogeneous first-order rate constant for microbial degradation $k = 0.1 \ day^{-1}$ (a 1-week half-life approx). $\Delta L = 1.00$ cm, to obtain the steady-state concentration profile of an additive to a penetration depth of 1 m. (Gale)
    Hint: Use the expression $A(N) = A(N - 1)\exp(-k\Delta T)$ to compute each successive element. Plot the percent injection (y-axis) vs. penetration depth (cm) (x-axis). The first, second, and third elements are 97.5, 95.1, and 92.8 ppm.

12. In a chlorine cell, a mercury cathode and a graphite anode are used in a solvent S at 25 °C: (Hg)Na(m = 8×10^{-3}) in $S_{(l)}$/NaCl in S/Cl$_{2(g)}$($p = 0.95$ atm)/C. The electrolyte is a salt dissolved in liquid S in a concentration of $31 \ g \ dm^{-3}$ at the same temperature. Each one of the electrodes has an area of $1.00 \ m^{-2}$ and the cell has a total resistance of 0.2 m2. The normal potential for the cathodic reaction is

$$Na^+ + e^- + xHg \rightleftharpoons Na(Hg)_x \qquad E^\circ_{Na^+/NaHg} = -1.868 \ V \text{ at } 25 \ °C$$

and for the anodic reaction:

$$Cl_2 + 2e^- \rightleftharpoons 2Cl^- \qquad E^\circ_{Cl^-/Cl_2} = 1.361 \ V \text{ at } 25 \ °C$$

The cathodic reaction takes place with an activation overpotential of -0.021 V, and the anodic reaction takes place according to the Tafel coefficients $a = 0.06$ V and $b = 0.18 \ V \ dec^{-1}$ for the current density expressed in amperes m^{-2}. Both processes occur without any mass transfer with an efficiency of 100% in each electrode. (a) Determine the open-circuit potential of the system. (b) Determine the potential difference needed to obtain a production of Cl$_2$ of $2.5 \ g \ hr^{-1}$. (c) If the overpotential for the evolution of H$_2$ on the cathode in the solvent S is -1.25 V at the same temperature, determine the optimum pH needed to suppress the hydrogen evolution at the same conditions. Consider that the activity coefficients for the ions are unity, and that the salt concentration does not vary during the process. Consider $A_{Cl} = 35.45 \ g \ mol^{-1}$ and $A_{Na} = 22.99 \ g \ mol^{-1}$. (Zinola)

MICRO RESEARCH PROBLEMS

1. If solar energy is to be the origin of the clean fuel, hydrogen, then there are two main ways in which this can be achieved. It is possible to use solar light on photovoltaic couples in air and to utilize the electricity that is produced to electrolyze water. Correspondingly, it is possible to directly irradiate semiconductor electrodes in electrochemical cells and thereby achieve the photoproduction of hydrogen and oxygen (photosplitting of water) directly without the need for two plants. Consider, first, a photoelectrochemical cell containing an n-type semiconductor as the photoanode and a Pt electrode as the cathode.

   (a) Knowing that the thermodynamic potential for the decomposition of water is 1.23 V, and taking i_0 for the cathodic reaction as 10^{-3} A cm^{-2} and i_0 for the anodic reaction as 10^{-8} A cm^{-2}, what outside potential would be needed to drive the cell (imagined to have a negligible IR drop) at 0.1 A cm^{-2}? (b) What energy gap in the photoanode would be necessary to drive the cell at this current density, assuming sufficient light could be brought to bear using concentrators? (c) Using Fig. 10.11, the solar spectrum, calculate what fraction of solar light would have photons of this energy. (d) What would be the maximum current density at which such a cell could be driven without concentrators? (e) Suppose that in another photochemical cell, the p-type cathode has an E_g of 1.1 eV and the photoanode, an E_g of 1.7 eV. At what maximum current density could this water splitter operate? (f) If it were necessary to use as the anode an n-type semiconductor that is not an oxide, what can be done to prevent its oxidation during operation? A fully quantitative treatment is required. (Bockris)

2. During the late 1990s the Daimler-Benz Company, in cooperation with the Ballard Company of Vancouver, Canada, stated that by 2003 it would mass produce electric cars driven from fuel cell electricity, the cells running on H_2 coming from the on-board re-forming of methanol or gasoline. (a) Examine quantitatively a scheme in which methanol is made chemically by the catalyzed chemical reaction, $CO_2 + 3H_2 \rightarrow CH_3OH + H_2O$.

   Assume the typical gas station refuels one car every 5 min for a 10-hr day and that the average car requires 15 gallons of gasoline. (b) Convert this to kilojoules per day and eventually to a (levelized) number of kilowatts supplied throughout a 24-hr day.

   Suppose that this energy is to be supplied to cars in the form of hydrogen in a fuel cell from re-formed methanol. (c) How many moles of CO_2 and H_2 would be required to form the methanol? (d) Calculate the cubic meters of air per day that would have to be collected at a given gas station to allow extraction of the necessary CO_2 (0.3% in air). (e) Then calculate the number of moles per day of H_2 needed to form the corresponding methanol (fuel cell efficiency is 60%). (f) Calculate the area of solar panels (number of square meters for a 10-kW sunny day) required to produce this H_2 at 20% efficiency for the conversion of solar

light to electricity and 80% efficiency for the electrolysis.

(g) Assuming that the area of the gas station is 15 meters squared (i.e. 15 × 15 m^2), examine whether it is possible to attain this energy by mounting solar panels within this area on a mast 500 m high. (h) Calculate, correspondingly, what surrounding area of houses would have to have solar panels on their roofs to supply one gas station with enough solar electricity to form the H_2 for the necessary methanol. (A sufficient supply of water is to be assumed.) (Bockris)

This page intentionally left blank

INDEX