CORROSION AND SURFACE CHEMISTRY OF METALS

© 2007, first edition, EPFL Press

Engineering Sciences *Materials*

CORROSION AND SURFACE CHEMISTRY OF METALS

Dieter Landolt

EPFL Press
A Swiss academic publisher distributed by CRC Press

© 2007, first edition, EPFL Press

CRC Press
Taylor & Francis Group

Taylor and Francis Group, LLC
6000 Broken Sound Parkway, NW, Suite 300,
Boca Raton, FL 33487

Distribution and Customer Service
orders@crcpress.com

www.crcpress.com

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress.

is an imprint owned by Presses polytechniques et universitaires romandes, a Swiss academic publishing company whose main purpose is to publish the teaching and research works of the Ecole polytechnique fédérale de Lausanne.

Presses polytechniques et universitaires romandes
EPFL – Centre Midi
Post office box 119
CH-1015 Lausanne, Switzerland
E-Mail: ppur@epfl.ch
Phone: 021/693 21 30
Fax: 021/693 40 27

www.epflpress.org

Originally published in French as *Corrosion et chimie de surfaces des métaux*
Copyright © 1993, 1997, 2003 Presses polytechniques et universitaires romandes

Translated and updated from the revised second French edition
© 2007, first edition, EPFL Press
ISBN 978-2-940222-11-7 (EPFL Press)
ISBN 978-0-8493-8233-8 (CRC Press)

Printed in Italy

All right reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprint, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publisher.

PREFACE

Throughout history, metals have had a decisive effect on human civilization, and they largely condition the way we live today. Electric power, combustion engines, ships, cars, railways, airplanes, production of fertilizers and pharmaceuticals, precision tools, machines, high rise buildings, sanitary systems, home appliances and even computers would not exist without metals. Corrosion and wear limit the useful lifetime of this equipment, reduce their efficiency and often affect their safety. Corrosion control and surface technology are therefore of crucial importance for our technology-based society.

Metals are found in nature in their oxidized state, mostly as oxide or sulfide minerals. The elaboration of metals from minerals requires energy, and, as a consequence, if left alone metals spontaneously return to their natural state; they corrode. The rate at which this occurs varies widely among different metals and as a function of their chemical environment, the prevailing temperature and the concurrent action of mechanical forces. Furthermore, the local corrosion rate of a metallic structure may be affected by electrochemical interactions between different parts of it and by its shape. Often, corrosion rates vary with time due to changing environmental conditions or because corrosion products accumulate on the surface. Corrosion resistance therefore is not an intrinsic property of a material but depends on a given metal-environment system taken as a whole. The prediction of corrosion behavior and the implementation of efficient corrosion control measures requires a good understanding of the underlying reaction mechanisms.

In past years our knowledge about metal surfaces and corrosion mechanisms has advanced greatly, and considerable progress has been achieved in controlling corrosion. Successful examples include multi-year warranties against perforation of car body panels, oil drilling in highly corrosive marine environments or safely operated power plants. Today, at the beginning of the twenty first century, corrosion science and engineering faces many new challenges. First of all, the evident need for sustainability of future development requires that wasting of materials and energy resources due to corrosion be reduced further by developing appropriate technologies. A number of well-established methods of corrosion protection need to be replaced by

© 2007, first edition, EPFL Press

more environment-friendly techniques. The development of ever smaller electronic and micromechanical devices and the emerging field of nanotechnology require new approaches to corrosion and wear control on an increasingly finer scale. On the opposite side, installations for energy production, civil engineering structures and transport systems tend to become ever larger and more complex, increasing the requirements for reliable risk analysis and lifetime prediction in relation to corrosion. Corrosion issues are also crucial for safe long-term storage of nuclear waste. Complex physical and chemical interactions between metal surfaces and biological systems govern the lifetime of medical implants and determine the extent to which microbial corrosion damages process equipment and water distribution systems. On a different level, corrosion based chemical and electrochemical shaping and surface treatment processes offer new opportunities for surface engineering and for the fabrication of both microsystems and of new types of nanostructures. To successfully tackle these and other challenges, engineers and scientists need a sound education in the basics of corrosion, electrochemistry and surface chemistry of metals.

The present book is the result of courses that the author has been teaching for many years, mostly to students of materials science and chemical engineering. The original French edition of the book was well received, requiring a second and third printing. For the present English edition the content has been enriched and updated, taking into account recent scientific and technical developments. Among these changes, more space has been given to advanced methods for the characterization of surfaces and interfaces, some sections such as those dealing with dealloying, pitting and tribocorrosion have been rewritten, and a new section on the prevention of microbially influenced corrosion has been added.

This being an introductory textbook, the understanding of principles, rather than technological detail, is emphasized throughout. A careful presentation of the electrochemistry of metals and oxide films is given, and the fundamentals and methods of surface chemistry relevant for corrosion are introduced. Whenever possible, simple quantitative models are used to promote the physical understanding of corrosion and protection phenomena. Because students often experience difficulties in bridging theoretical notions and engineering applications, the practical relevance of theoretical concepts is repeatedly stressed and an entire chapter is dedicated to the presentation of commonly used approaches to corrosion control. Thus the book provides a comprehensive introduction to the principles and methods of modern corrosion science and engineering.

Senior or graduate students in materials science, chemistry and chemical engineering as well as professional scientists and engineers faced with problems of corrosion and surface engineering should profit from this book. It can be used as a textbook in university level courses or for self study. The content is suited for general corrosion courses as well as for introductory courses on the electrochemistry and surface chemistry of metals and oxides. Indeed, the selection of material can be easily adapted to different needs. For example, a corrosion course for engineering students may introduce just the most basic notions of Chapters 2, 4, and 6, with more

© 2007, first edition, EPFL Press

emphasis on the practical aspects of corrosion mechanisms of Chapters 7 to 11 and on corrosion prevention discussed in Chapter 12. A science oriented course for students in materials, chemistry or applied physics will focus primarily on the fundamentals and methods exposed in Chapters 2 to 6 and in parts of Chapters 8 and 12. Advanced corrosion courses may study in more detail such topics as high temperature corrosion, stress corrosion cracking and environmental effects on friction and wear presented in Chapters 9 to 11. Homework problems for each chapter complement the material treated in the text.

Dr. Fred Fenter of PPUR carefully translated the original French text which served as a basis for this book, and he was in charge of the production process. The author expresses his gratefulness and thanks to Fred Fenter for encouraging the preparation of this book and for the continuing support and precious collaboration throughout its realization.

Lausanne, December 2006
Dieter Landolt

© 2007, first edition, EPFL Press

TABLE OF CONTENTS

PREFACE v

Chapter 1 INTRODUCTION 1
1.1 The Corrosion of Metals 1
1.2 Basic Concepts 3
1.3 Phenomenological Aspects of Corrosive Attacks 9

Chapter 2 THERMODYNAMICS OF CORROSION REACTIONS 15
2.1 Chemical Equilibrium 15
2.2 Electrochemical Equilibrium 22
2.3 Equilibrium Potential and Corrosion Behavior 31
2.4 A Closer Look at Electrochemical Cells 47
2.5 Physical Potential Scale 54

Chapter 3 SURFACES AND INTERFACES 59
3.1 Surface Tension Effects 59
3.2 Adsorption 66
3.3 Chemical Characterization of Surfaces 73
3.4 Structural Characterization of Surfaces 89
3.5 Electrode-Electrolyte Interface 98

Chapter 4 CORROSION REACTION RATES 119
4.1 Basic Concepts 119
4.2 Charge Transfer 125
4.3 Mass Transport 142

Chapter 5 EXPERIMENTAL STUDY OF ELECTRODE REACTIONS 179
5.1 Electrochemical Kinetics and Reaction Mechanisms 179
5.2 Electrochemical Non-Steady-State Methods 192

Chapter 6 PASSIVITY OF METALS 227
6.1 Transition from Active to Passive State 227
6.2 Growth and Properties of Passive Oxide Films 240
6.3 Passivity Breakdown and Transpassive Behavior 262

© 2007, first edition, EPFL Press

Chapter 7 LOCALIZED CORROSION PHENOMENA 275
7.1 Corrosion Cells 275
7.2 Non-Uniform Corrosion Governed by Metallurgy 291
7.3 Pitting Corrosion 312

Chapter 8 OXIDATION AND ATMOSPHERIC CORROSION 331
8.1 Low-Temperature Oxidation 331
8.2 Atmospheric Corrosion of Steel 343
8.3 Atmospheric Corrosion of Non-Ferrous Metals 360

Chapter 9 HIGH-TEMPERATURE CORROSION 365
9.1 Basic Concepts 365
9.2 Oxidation of Metals 372
9.3 Oxidation of Alloys 388
9.4 Corrosion in the Presence of Sulfur Compounds 402

Chapter 10 TRIBOCORROSION 415
10.1 Introduction 415
10.2 Chemical Effects in Friction and Lubrication 417
10.3 Wear and Corrosion in Sliding Contacts 428
10.4 Wear and Corrosion Due to Impacts 444

Chapter 11 ENVIRONMENT INDUCED CRACKING 461
11.1 Introduction 461
11.2 Experimental Study of Environment Induced Cracking 465
11.3 Hydrogen Embrittlement 477
11.4 Stress Corrosion Cracking 491
11.5 Corrosion Fatigue 502

Chapter 12 PROTECTION OF ENGINEERING SYSTEMS AGAINST CORROSION 513
12.1 Preventive Measures 513
12.2 Coatings 524
12.3 Corrosion Inhibitors 544
12.4 Electrochemical Protection 563

ANNEXES 585
A.1 Constants and Units 585
A.2 Periodic Table of the Elements 587
A.3 Problems 588
A.4 List of Symbols 611

© 2007, first edition, EPFL Press

CHAPTER 1

INTRODUCTION

1.1 THE CORROSION OF METALS

1.1.1 Overview of the phenomena

Corrosion, from the Latin *corrodere*, means "to chew away", "to attack". It is estimated that corrosion destroys one quarter of the world's annual steel production, which corresponds to about 150 million tons per year, or 5 tons per second. Of course, corrosion is not limited to steel but affects all materials: metals, polymers and ceramics. It is the result of chemical and/or physical interactions between the material and its environment. Examples of the corrosion phenomena include:

- transformation of steel into rust;
- cracking of brass in the presence of ammonia;
- oxidation of an electrical contact made of copper;
- weakening of high-resistance steal by hydrogen;
- hot corrosion of a super-alloy in a gas turbine;
- swelling of PVC in contact with a solvent;
- chemical attack of a nylon tube by an oxidizing acid;
- alkaline attack on refractory bricks;
- chemical attack of mineral glass by an alkaline solution.

Metals differ from other materials by a number of favorable properties: ductility, high tensile strength, temperature resistance, electrical and thermal conductivity and ease of joining and machining. Critical elements of machines, airplanes, cars, electrical power plants, precision instruments, civil engineering structures and chemical plants are normally made of metal. Electronic components and devices also contain numerous metallic elements to provide electrical connections. Quite generally, the durability and life time of installations, machines and devices is critically dependent on their corrosion and wear resistance. Because of their unique position among engineering materials the study of the corrosion and protection of metals is an important part of materials science and engineering.

Interestingly, most metals and alloys are not thermodynamically stable in contact with the atmosphere and with water and they should spontaneously corrode. Fortunately, for most applications the rate of corrosion can be kept sufficiently small by using adequate preventive measures (choice of materials, surface treatment,

© 2007, first edition, EPFL Press

electrochemical protection, etc.). In this way metallic objects can satisfactorily fulfill their function over their projected lifetime.

Some definitions

From the point of view of the construction engineer ***corrosion*** is damaging: it destroys a material or degrades its functional properties, rendering it unsuitable for the intended use. ***Corrosion damage*** is the degradation of a material or of its functional properties through a chemical reaction with the environment. Materials can also be damaged by ***wear***, which results from rubbing between solid surfaces or from impingement of fluids or of solid particles. Wear causes a progressive loss of material from a surface by mechanical mechanisms, but chemical interactions between the material and its environment often slow or accelerate the damage. The study of corrosion and protection of metals must also include degradations that arise as a result of combined mechanical and chemical effects.

Sometimes, corrosion is a welcome, even desirable, phenomenon. For example, corrosion destroys metallic objects abandoned in nature and thus eliminates them. Corrosion reactions are also used in industrial manufacturing. A well-known example is the anodizing of aluminum where one applies an anodic voltage to the metal in a suitable electrolyte. Anodizing reinforces the natural oxide film at the surface and thus provides improved corrosion resistance and sometimes also a decorative effect. Similarly, in chemical and electrochemical polishing corrosion reactions are used to produce a smooth surface finish. We can therefore define ***corrosion*** in a general way as follows [1]: *Corrosion is an irreversible interfacial reaction of a material with its environment, resulting in the loss of material or in the dissolving of one of the constituents of the environment into the material*. This definition includes both the positive and negative effects of corrosion. It also includes material damage due to the absorption of a constituent of the environment, such as hydrogen absorption into steel, which causes embrittlement and thus impairs the mechanical properties of the material.

1.1.2 The economic importance of the corrosion

Corrosion affects all areas of the economy, from the integrated circuit to the bridge made of reinforced concrete. The cost of corrosion has been estimated to represent 4% of the gross national product [2]. Even for a small country like Switzerland, this number represents several billion Euros per year. These numbers include:

- direct losses: replacement of corroded materials and of equipment ruined by corrosion;
- indirect losses: cost of repair and loss of production;
- cost of corrosion protection: use of more expensive corrosion-resistant materials, application of surface coatings, cathodic protection systems ;
- cost of corrosion prevention: maintenance, inspections, corrosion prevention by design.

© 2007, first edition, EPFL Press

The direct losses represent only part of the total costs of corrosion and are, in fact, often inferior to the indirect costs. If a nuclear power plant, which represents a significant capital investment, has to be stopped for repair or replacement of a corroded heat exchanger, the price of the part is insignificant relative to the cost of the lost production time. Similarly, in order to replace a corroded hot water pipe, buried in the wall of a building, the cost of the repair will usually largely exceed the price of the pipe itself.

The many different types of expenses involved make estimates of the total cost of corrosion difficult and uncertain. There is no doubt, however, that the amounts of money involved are quite elevated. Another important aspect of corrosion concerns safety. Corrosion damage can impair the safe operation of installations or machines and be at the origin of severe accidents and the loss of human life. Furthermore, corrosion is a waste of raw materials and, indirectly, of energy.

1.2 BASIC CONCEPTS

1.2.1 Corrosion reactions

With only a few exceptions (c.f. Chapter 10), the corrosion of metals is due to an irreversible oxidation-reduction (redox) reaction between the metal and an oxidizing agent present in the environment. The oxidation of the metal is inseparably coupled to the reduction of the oxidizing agent:

metal + oxidizing agent → oxidized metal + reducing agent

For example, the corrosion of iron in the presence of hydrochloric acid (Figure 1.1) is due to reaction (1.1).

$$Fe_{(s)} + 2\,HCl_{(aq)} \rightarrow FeCl_{2(aq)} + H_{2(g)} \tag{1.1}$$

In this equation, the indices (s), (aq), and (g) refer to the solid, aqueous and gaseous phases, respectively. In aqueous phase, hydrochloric acid and ferrous chloride exist in ionic form. We can therefore also write:

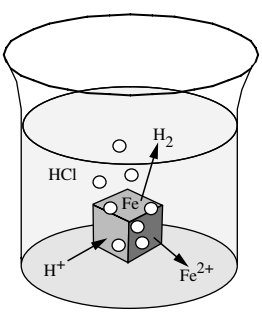


Figure 1.1 The corrosion of iron in acidic solution.

© 2007, first edition, EPFL Press

$$Fe + 2\,H^{+}_{(aq)} + 2\,Cl^{-}_{(aq)} \rightarrow Fe^{2+}_{(aq)} + 2\,Cl^{-}_{(aq)} + H_{2(g)} \tag{1.2}$$

In this case, the oxidizing agent is the solvated proton, $H^{+}_{(aq)}$. The products of the reaction are the solvated ferrous ion $Fe^{2+}_{(aq)}$, and gaseous hydrogen $H_{2(g)}$. To keep the equations simple, the indices that indicate the phase are usually not shown. Because the chloride ions do not undergo a chemical change, the reaction (1.2) can also be expressed as follows.

$$Fe + 2H^{+} \rightarrow Fe^{2+} + H_2 \tag{1.3}$$

Under neutral and alkaline conditions, the corrosion of metals is generally due to a reaction of the metal with oxygen. For example, when exposed to air and to humidity iron forms ***rust***, FeOOH.

$$4\,Fe + 3\,O_2 + 2\,H_2O \rightarrow 4\,FeOOH \tag{1.4}$$

Metals also react with oxygen in acidic environments, but under these conditions the concentration of oxygen is usually so much lower than that of protons that its effect can be neglected.

In the absence of humidity, gaseous oxygen is a corrosive agent only at elevated temperatures (at several hundred degrees Celsius). For this reason, a distinction is made between *wet corrosion*, or corrosion at ambient temperatures, and *dry corrosion*, or corrosion at elevated temperatures.

Oxidizing Agents

In wet corrosion, two principal oxidizing agents are encountered in practice:

- solvated protons;
- dissolved oxygen.

However, other oxidizing agents can also cause corrosion in wet environments, such as:

- oxidizing metal cations: Cu^{2+}, Fe^{3+}, Sn^{4+};
- oxidizing anions: NO_2^- , NO_3^-, CrO_4^{2-}, MnO_4^-, OCl^-;
- dissolved oxidizing gases: O_3, Cl_2, SO_3.

At high temperatures, certain chemical substances, normally inoffensive, become corrosive. Among the oxidizing agents responsible for dry corrosion, one finds:

- gaseous oxygen;
- water vapor;
- carbon dioxide, CO_2;
- sulfur-containing compounds: S_2, SO_2, SO_4^{2-}.

Partial reactions

All redox reactions consist of two partial reactions, also sometimes called half-cell reactions: the partial oxidation reaction, also referred to as the ***anodic partial***

© 2007, first edition, EPFL Press

reaction; and the partial reduction reaction or ***cathodic partial reaction***. Thus for reaction (1.3):

$$\mathrm{Fe} \rightarrow \mathrm{Fe}^{2+} + 2\,e \qquad \text{anodic partial reaction} \tag{1.5}$$

$$2\,\mathrm{H}^{+} + 2\,e \rightarrow \mathrm{H}_2 \qquad \text{cathodic partial reaction} \tag{1.6}$$

$$\mathrm{Fe} + 2\,\mathrm{H}^{+} \rightarrow \mathrm{Fe}^{2+} + \mathrm{H}_2 \qquad \text{overall reaction} \tag{1.3}$$

The anodic and cathodic partial reactions show explicitly the electrons that are exchanged during the course of the redox reaction.

Any chemical transformation that implies the transfer of charge across the interface between an electronic conductor (the electrode) and an ionic conductor (the electrolyte) is referred to as an ***electrochemical reaction***. An electrochemical reaction can include one or several ***electrode reactions***. For example the reaction (1.3) is an electrochemical reaction: each atom of iron that passes into solution implies the exchange of two electrons between the metal and the protons. Two electrode reactions are involved: the oxidation of the iron and the reduction of the proton. According to the definition given above, *all corrosion reactions that involve metal oxidation are electrochemical reactions*. In order to understand and control corrosion phenomena it is essential to study the thermodynamics and kinetics of electrochemical reactions.

1.2.2 Faraday's Law

According to Faraday's Law, when n_i moles of a given substance react, a proportional electric charge Q passes across the electrode-electrolyte interface:

$$Q = n \cdot F \cdot n_\mathrm{i} \tag{1.7}$$

Here, F designates a universal constant, called ***Faraday constant***: $F = 96{,}485$ C/mol. The ***charge number*** n, without units, expresses the stoichiometric coefficient of the electrons in the equation for the electrode reaction. For example, for the reaction (1.5) describing the anodic dissolution of iron, $n = 2$.

By taking the derivative of equation (1.7) with respect to time, one obtains Faraday's law in the form it is most often used:

$$I = n\,F\,\frac{\mathrm{d}n_\mathrm{i}}{\mathrm{d}t} \tag{1.8}$$

where the electric current, in amperes, is given by $I = \mathrm{d}Q/\mathrm{d}t$ and $\mathrm{d}n_\mathrm{i}/\mathrm{d}t$ represents the rate of reaction in moles per second. Faraday's law thus states that the rate of an electrode reaction is proportional to the magnitude of the electrical current that crosses the electrode-electrolyte interface.

An ***electrochemical cell*** is a system composed of two electrodes in contact with an electrolyte and connected by an electric conductor. Corrosion cells, batteries and fuel cells are examples of electrochemical cells. As an illustration, Figure 1.2 schematically shows the functioning of an H_2/O_2 fuel cell. At the anode, hydrogen is

© 2007, first edition, EPFL Press

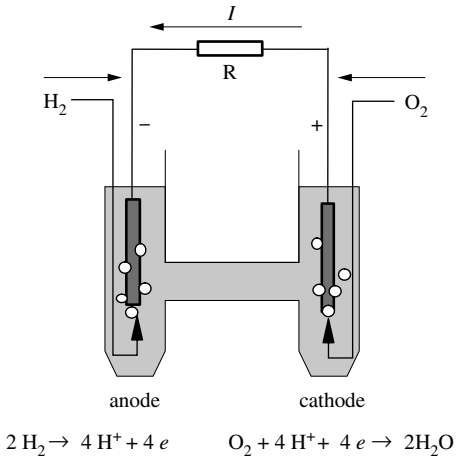


Figure 1.2 H_2/ O_2 fuel cell.

oxidized to H^+, whereas at the cathode, oxygen is reduced to H_2O. The overall reaction corresponds to an oxidation-reduction reaction between oxygen and hydrogen.

$$\text{anode:} \quad 2\,H_2 \rightarrow 4\,H^+ + 4\,e \tag{1.9}$$

$$\text{cathode:} \quad O_2 + 4\,H^+ + 4\,e \rightarrow 2\,H_2O \tag{1.10}$$

$$\text{overall reaction:} \quad O_2 + 2\,H_2 \rightarrow 2\,H_2O \tag{1.11}$$

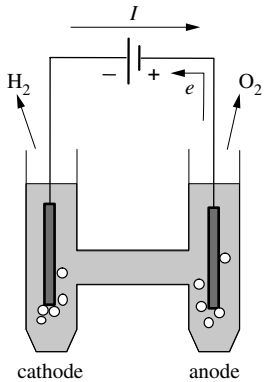
The H_2/O_2 fuel cell uses inert electrodes (porous graphite containing platinum as a catalyst, for example). A solution of phosphoric acid, or some other acid, serves as electrolyte. The reaction between oxygen and hydrogen is spontaneous and the fuel cell therefore provides electric power. The current is directly proportional to the number of moles of H_2 and O_2 that react at the electrodes, with $n = 4$ and $n = 2$ for oxygen and hydrogen, respectively:

$$I = 4\,F\,\frac{\mathrm{d}n_{O_2}}{\mathrm{d}t} = 2\,F\,\frac{\mathrm{d}n_{H_2}}{\mathrm{d}t} \tag{1.12}$$

In an ***electrolysis cell,*** also called ***electrolyzer***, an external voltage (or current) is applied to the electrodes in order to carry out an electrolysis. For example, in a water electrolysis cell, the current which passes between the anode and the cathode decomposes water into oxygen and hydrogen (Figure 1.3). The number of moles of H_2 and O_2 produced during electrolysis is also given by Faraday's law.

In an electrolysis cell, the anode is the positive pole, whereas in a fuel cell or battery it is the negative one. In all cases, however, the partial oxidation reaction takes place at the anode and the partial reduction reaction at the cathode. In other words, the anode is always the site of oxidation, independent of its polarity.

© 2007, first edition, EPFL Press



$4\,H^+ + 4\,e \rightarrow 2\,H_2 \qquad 2\,H_2O \rightarrow O_2 + 4\,H^+ + 4\,e$

Figure 1.3 Electrolysis of water.

1.2.3 Corrosion rate

The rate of corrosion can be expressed in many different ways depending on the application and personal preferences:

- as a mass loss per unit of surface and per unit of time;
- as the number of moles transformed per unit of surface and per unit of time;
- as corroded depth per unit of time;
- as a current density.

Table 1.4 indicates the conversion factors between units that are frequently used for measuring corrosion rate.

Table 1.4 Frequently used units for measuring corrosion rate. To obtain the units indicated in the first row, one multiplies the unit in the first column by the corresponding conversion factor.

	$\frac{\text{mole}}{\text{m}^2\text{s}}$	$\frac{\text{mole}}{\text{cm}^2\text{s}}$	$\frac{\text{A}}{\text{m}^2}$	$\frac{\mu\text{A}}{\text{cm}^2}$	$\frac{\text{mg}}{\text{dm}^2\text{day}}$	$\frac{\text{mm}}{\text{year}}$
$\frac{\text{mole}}{\text{m}^2\text{s}}$	1	10^{-4}	$9.65\times10^4\,n$	$9.65\times10^6\,n$	$8.64\times10^5\,M$	$3.15\times10^4\,\frac{M}{\rho}$
$\frac{\text{mole}}{\text{cm}^2\text{s}}$	10^4	1	$9.65\times10^8\,n$	$9.65\times10^{10}\,n$	$8.64\times10^9\,M$	$3.15\times10^8\,\frac{M}{\rho}$
$\frac{\text{A}}{\text{m}^2}$	$\frac{1.04\times10^{-5}}{n}$	$\frac{1.04\times10^{-9}}{n}$	1	100	$8.96\,\frac{M}{n}$	$0.327\,\frac{M}{n\rho}$
$\frac{\mu\text{A}}{\text{cm}^2}$	$\frac{1.04\times10^{-7}}{n}$	$\frac{1.04\times10^{-11}}{n}$	0.01	1	$8.96\times10^{-2}\,\frac{M}{n}$	$3.27\times10^{-3}\,\frac{M}{n\rho}$
$\frac{\text{mg}}{\text{dm}^2\text{day}}$	$\frac{1.16\times10^{-6}}{M}$	$\frac{1.16\times10^{-10}}{M}$	$0.112\,\frac{n}{M}$	$11.2\,\frac{n}{M}$	1	$\frac{3.65\times10^{-2}}{\rho}$
$\frac{\text{mm}}{\text{year}}$	$3.17\times10^{-5}\,\frac{\rho}{M}$	$3.17\times10^{-9}\,\frac{\rho}{M}$	$3.06\,\frac{n\rho}{M}$	$306\,\frac{n\rho}{M}$	$27.4\,\rho$	1

M = atomic mass in g/mol, ρ = volumetric mass en g/cm^3, n = charge number (dimensionless).

© 2007, first edition, EPFL Press

Influence of potential

The prevailing electrode potential determines the direction and rate of the electrode reactions. For example, in a 1 molar solution of Fe^{2+} and Fe^{3+} ions that is in contact with a platinum electrode, the electrode reaction (1.13) takes place.

$$Fe^{3+} + e = Fe^{2+} \tag{1.13}$$

At an electrode potential of about 0.77 V (c.f. Chapter 2) the above reaction is at equilibrium. The reaction goes from the left to the right if the potential is inferior to this value, and from the right to the left if the potential is more positive. Corrosion reactions take place in an analogous fashion. In hydrochloric acid, for example, iron dissolves less rapidly if one shifts the potential in the negative direction. This effect is used in practice for the protecting metals against corrosion by imposing a sufficiently negative potential.

Influence of corrosion products

The physical nature of corrosion products often has a major influence on the rate of corrosion. We can divide products into three basic types:

- dissolved ions;
- porous films;
- compact films.

The corrosion of metals in an acidic environment normally produces hydrated cations or complexes that dissolve in the electrolyte and diffuse away from the metal surface. Most often they do not affect the rate of corrosion in an appreciable way.

In neutral or basic environments, but sometimes even in acid solutions, corrosion products are often only slightly soluble. They therefore precipitate at the surface of the metal in the form of hydroxides or metal salts, forming a porous or non-compact surface film. Such films do not protect the metal from corrosion, although they may slow its rate. Rust formed on the surface of steel exposed to a humid atmosphere is an example of this type of corrosion product.

Compact films, in particular oxide films, form a barrier between the metal and its environment, thus protecting the metal. The corrosion of iron in a dry atmosphere stops after the formation of an oxide film that is only a few nanometers thick. On certain metals, compact films also form in humid or liquid environments. These are referred to as ***passive films***. They are present on stainless steel in aqueous environments, on aluminum exposed to humid air, and on carbon steel in alkaline surroundings, to cite a few examples. The rate of corrosion of metals coated with a passive film usually is very low.

At high temperatures, under dry corrosion conditions, cations and anions can diffuse more easily through a compact oxide film. The film therefore grows progressively to a greater thickness and this corresponds to a corrosion reaction. Figure 1.5 shows schematically the reaction of a metal with oxygen under these conditions. The metal cations produced by the oxidation reaction at the oxide-metal interface diffuse towards the exterior of the film. At the outer surface, oxygen is

© 2007, first edition, EPFL Press

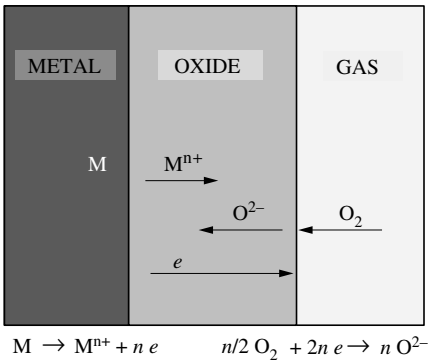


Figure 1.5 High-temperature corrosion: reaction between the metal and oxygen in the presence of an oxide layer having both ionic and electronic conductivity.

reduced to O^{2-} anions that diffuse towards the oxide-metal interface. The electrons that are liberated at the oxide-metal interface cross the oxide film and react at the oxide-gas interface. The outer surface of the oxide film can therefore be thought of as the cathode, and the inner surface as the anode. In contrast to classical electrochemical cells, the oxide here is both electrolyte and electron conductor.

The different examples illustrate how the rate of corrosion reactions can be affected by the properties of the corrosion products. Solid corrosion products often form slowly and sometimes they may change their properties with time. This can present a problem when accelerated laboratory tests are used for the prediction of the long-term performance of a metal. In order to extrapolate this type of results, it is necessary to well understand the reaction mechanisms involved.

1.3 PHENOMENOLOGICAL ASPECTS OF CORROSIVE ATTACKS

1.3.1 Uniform and localized corrosion

The reaction shown in Figure 1.1 occurs uniformly over the entire iron surface. On the macroscopic scale, it is not possible to distinguish between the cathodic and anodic sites–the entire surface of the metal acts simultaneously as cathode and anode. This type of behavior is referred to as ***uniform corrosion***.

It is possible to separate the cathode from the anode by putting the iron in contact with a platinum electrode, thus creating an electrochemical cell (Figure 1.6). The cathodic partial reaction (1.6) takes place more easily on the platinum, whose surface acts as a catalyst, than on the iron surface. As a result, a significant production of hydrogen is observed on the platinum, while at the hydrogen production on the iron practically stops. Simultaneously, the corrosion rate of iron increases. In this example, the anodic partial reaction takes place exclusively on the iron, whereas the cathodic partial reaction, which here limits the corrosion rate, takes place mostly on the

© 2007, first edition, EPFL Press

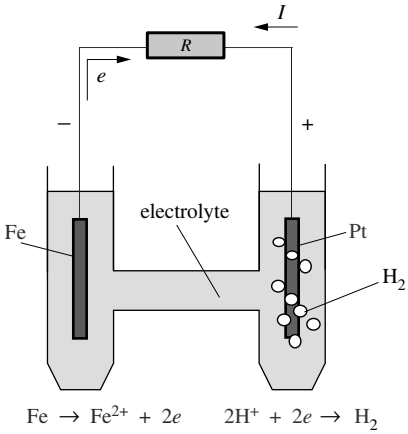


Figure 1.6 Iron-platinum corrosion cell in an acid electrolyte: dissolution of iron at the anode and reduction of protons at the platinum cathode.

platinum. The electrons that are liberated at the iron anode travel through the external conductor towards the cathode, where they are consumed for the reduction of protons. The overall reaction is still given by reaction (1.3). In our example, iron and platinum form a ***corrosion cell*** in which iron acts as anode and platinum as cathode. Corrosion cells are at the origin of ***localized corrosion,*** meaning corrosion phenomena where the anodic and the cathodic sites on a metal surface or on a metallic object can be clearly distinguished (Figure 1.7).

In practice, localized corrosion can be due to heterogeneities either in the material or in the environment to which a material is exposed (see Chapter 7). Usually, localized corrosion is more difficult to control than uniform corrosion.

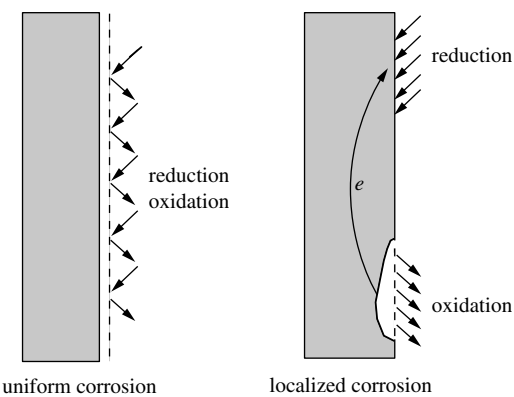


Figure 1.7 Uniform and localized corrosion of a metal surface.

© 2007, first edition, EPFL Press

1.3.2 Different types of corrosion damage and their origin

The corrosion behavior of engineering materials is influenced by many factors (Figure 1.8):

- chemical composition and microstructure of the metal;
- chemical composition of the environment;
- physical parameters such as temperature, convection, irradiation, etc.;
- mechanical forces including tensile stresses, shocks or friction.

Because the corrosion behavior of metals is governed by complex interactions involving many parameters, it manifests itself in numerous often unanticipated forms. The corrosion resistance of a given metal is not an intrinsic property of that metal, but a ***systems property***. The same metal may rapidly corrode in a certain environment while under different conditions it is stable. From a more fundamental point of view, the corrosion resistance of metals is essentially determined by the reactivity of the ***metal-environment interface.*** Therefore the chemical and structural characterization of surfaces and interfaces (cf. Chapter 3) and the study of their electrochemical behavior in a given environment (cf. Chapter 4) are important aspects of corrosion science.

Certain types of corrosion damages found in practice present a characteristic appearance that provides a hint as to their probable origin and mechanism. In his introduction to corrosion engineering (cf. the general bibliography at the end of this chapter), M. Fontana proposes the grouping of corrosion damage into eight types:

Uniform corrosion is a loss of material distributed uniformly over the entire surface exposed to the corrosive environment. Metals in contact with strong acids are sometimes subject to uniform corrosion.

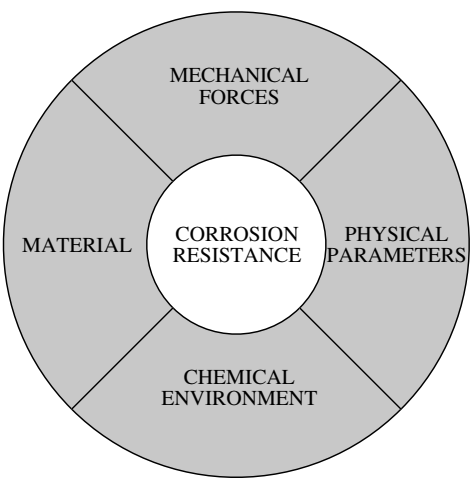


Figure 1.8 Corrosion system.

© 2007, first edition, EPFL Press

Galvanic corrosion, also called bimetallic corrosion, results from the formation of an electrochemical cell between two metals. The corrosion of the less noble metal is thus accelerated.

Crevice corrosion is caused by a difference of oxygen availability between two sites on a passive metal that leads to the formation of an electrochemical cell. A selective attack within cracks and at other sites of poor oxygen access is frequently observed.

Pitting corrosion is observed on passive metals in presence of certain anions (in particular chloride) when the potential exceeds a critical value. This process typically produces cavities with diameters on the order of several tens of micrometers.

Intergranular corrosion is a selective attack of grain boundaries. Often, it is related to thermal treatments that lead to preferred precipitation of phases at grain boundaries.

Selective corrosion, also called selective leaching or dealloying, implies the selective dissolution of one of the components of an alloy that forms a solid solution. It leads to the formation of a porous layer made of the more noble metal.

Erosion corrosion is the result of an electrochemical reaction combined with a material loss by mechanical wear due to impingement of solids or a fluid.

Stress corrosion cracking results from the combined action of corrosion and of mechanical stress. It manifests itself by crack formation at stress levels well below the ultimate tensile strength of a material.

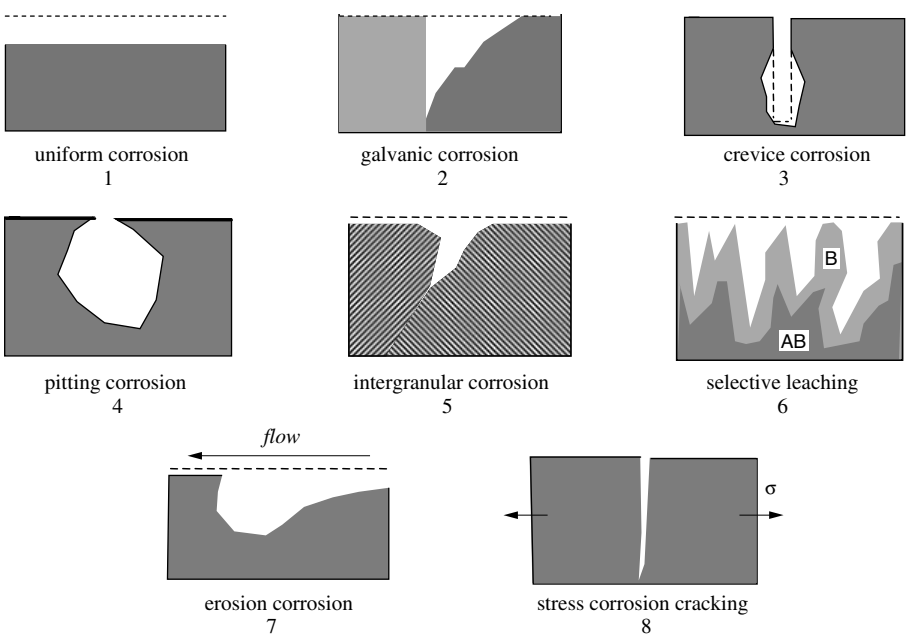


Figure 1.9 Eight types of corrosion damage.

© 2007, first edition, EPFL Press

Figure 1.9 schematically shows these different types of corrosion, all of which will be discussed in more detail in the following chapters, with an emphasis on the corresponding reaction mechanisms. We will see that this classification in eight categories does not adequately take into account all the fundamental mechanisms responsible for corrosion. However, this system is useful for the engineer in his or her attempt to identify the origin of certain types of damage due to corrosion.

General bibliography

M. Fontana, *Corrosion Engineering*, 3rd edition, McGraw Hill International Edition, New York, 1987, 556 pp.

H. H. Uhlig, R. W. Revie, *Corrosion and Corrosion Control*, 3rd edition, John Wiley, New York, 1985, 441 pp.

G. Wranglen, *An Introduction to Corrosion and Protection*, Chapman and Hall, London, 1985, 288 pp.

References

[1] K. E. Heusler, D. Landolt, S. Trasatti, *Pure and Appl. Chem.* 61, 19 (1989).

[2] R. M. Latanision, *Materials Performance* 26, 9 (1987).

© 2007, first edition, EPFL Press

CHAPTER 2

THERMODYNAMICS OF CORROSION REACTIONS

2.1 CHEMICAL EQUILIBRIUM

2.1.1 Basic concepts

Under which conditions is it possible for metal corrosion to occur? The answer to this fundamental question requires the study of chemical and electrochemical equilibria; only then can we begin to understand and to quantify corrosion phenomena. The chapter begins with some basic concepts about chemical equilibrium; these will then be applied specifically to corrosion reactions.

In thermodynamics, chemical reactions are often expressed in the following manner:

$$0 = \sum \nu_i \, B_i \tag{2.1}$$

where the stoichiometric coefficients ν_i of the reaction products are positive, and those of the reactants are negative. The symbol B_i designates the chemical species that participate in the reaction.

The ***Gibbs free energy of reaction*** ΔG is defined by

$$\Delta G = \sum \nu_i \, \mu_i \tag{2.2}$$

A species *i* has a ***chemical potential*** μ_i equal to its partial free energy which, in the case of a pure substance, is equivalent to the molar free energy.

The chemical potential of a species *i*, present in a mixture, can be separated into two parts:

$$\mu_i = \mu_i^\circ + RT \ln a_i \tag{2.3}$$

The ***standard chemical potential,*** μ_i°, depends only on the pressure and the temperature, while the logarithmic term includes the activity, a_i, which depends on the concentration as will be discussed later in this chapter.

We now introduce two new definitions:

- The standard free energy of reaction ΔG°;

© 2007, first edition, EPFL Press

$$\Delta G^\circ = \sum \nu_i \ \mu_i^\circ \tag{2.4}$$

- and the equilibrium constant K.

$$\ln K = \sum \nu_i \ \ln a_i = \ln \prod a_i^{\nu_i} \tag{2.5}$$

This allows us to re–express equation (2.2):

$$\Delta G = \Delta G^\circ + RT \ln K \tag{2.6}$$

Spontaneous reactions are associated with a decrease in the free energy, $\Delta G < 0$. Chemical equilibrium therefore corresponds to the minimum in free energy,

$$\Delta G = 0 \tag{2.7}$$

which in turn provides a relationship between the equilibrium constant K and the standard free energy of reaction ΔG°:

$$\Delta G^\circ = -RT \ \ln K \tag{2.8}$$

The ***activity*** is the product of an activity coefficient, f_i, and a concentration. Equation (2.3) defines the activity of a species i in a given mixture. At standard state, $\Delta G = \Delta G^\circ$, implying that $a_i = 1$. However, the choice of this standard state and of the units of concentration is arbitrary and established by the following conventions.

Convention (I)

This convention applies specifically to gases. The standard state is the perfect gas at $P_{o,i} = 1$ bar or $P_{o,i} = 1$ atm (1 atm = 1.013 bar). In a mixture, the activity of a gas i, also called the fugacity, has the following value:

$$a_i = f_i \frac{P_i}{P_{o,i}} \tag{2.9}$$

As elevated pressures are exceptional, in most practical situations the gases involved in corrosion reactions behave as perfect gases. We can therefore put $f_i = 1$, which gives as a result:

$$a_i = P_i \tag{2.10}$$

But keep in mind that a_i does not have units and, consequently the units of P_i must correspond to those of the standard state. In this book, for the standard state for gases we generally assume that 1 atm ≈ 1 bar.

Convention (II)

This convention applies to solvents, in particular water, and to the components of an alloy. The standard state is the species i in its pure state. Concentration is expressed in

© 2007, first edition, EPFL Press

molar fraction $X_i = N_i/\Sigma N_j$ where N_i represents the number of moles of i and ΣN_j the total number of moles.

$$a_i = f_i X_i \tag{2.11}$$

According to this convention, the activity coefficient $f_i = 1$ if $X_i = 1$.

Convention (III)

This convention applies to solutes, for example to salts dissolved in water. The standard state of species i is a solution of concentration $c_{o,i} = 1$, which has the properties of an infinitely dilute solution. This does not refer to a real physical situation, but rather offers a formalism that is quite useful for the calculation of equilibria in solution. The concentration can be expressed in weight percent, molarity (mole per liter), or molality (mole per kg).

$$a_i = f_i \frac{c_i}{c_{o,i}} \tag{2.12}$$

In general, instead of (2.12) one writes simply:

$$a_i = f_i c_i \tag{2.13}$$

In this case, the units of concentration, on which the values of f_i and a_i depend, must correspond to that of the standard state. In the study of aqueous corrosion and electrochemistry, the concentration of dissolved species is most often expressed in moles per liter.

2.1.2 Standard free energy of reaction

In order to calculate $\Delta G°$, it is necessary to know $\mu_i°$ (2.4). For this, we set $\mu_i° = \Delta G_i°$ where $\Delta G_i°$ represents the ***molar free energy of formation*** of the reactants and products, determined starting from the elements in their natural state (gas, liquid, solid), for the temperature of the reaction under consideration and at a pressure of 1 bar. These values can be found in reference [1].

The ***Gibbs–Helmholtz equation*** relates the standard free energy of reaction $\Delta G°$, the standard enthalpy of reaction $\Delta H°$ and the standard entropy of reaction $\Delta S°$:

$$\Delta G° = \Delta H° - T\,\Delta S° \tag{2.14}$$

Often, the assumption is made that $\Delta H°$ et $\Delta S°$ do not vary with temperature; in this case they can be replaced by the constants A and B.

$$\Delta G° = A + B\,T \tag{2.15}$$

Equation (2.15) is a good approximation of the true reaction behavior over a given range of temperatures. Table 2.1 provides the numerical values of A and B for a number of reactions useful in the study of corrosion.

© 2007, first edition, EPFL Press

Table 2.1 Standard free energies of several reactions (adapted from ref. [2]).

Reaction	$-A$ (kJ mol^{-1})	B (J mol^{-1} K^{-1})	Range of validity (°C)
$2\,Al + 3/2\,O_2 = Al_2O_3$	1673	313	25 – 659
$C + 1/2\,O_2 = CO$	112	– 87.8	25 – 2000
$C + O_2 = CO_2$	394	– 1.13	25 – 2000
$C + 2H_2 = CH_4$	90.2	377	25 – 2000
$Co + 1/2\,O_2 = CoO$	240	78.0	25–1700
$2\,Cr + 3/2\,O_2 = Cr_2O_3$	1135	257	25–1898
$23\,Cr + 6\,C = Cr_{23}C_6$	411	– 38.7	25 – 1400
$Fe + 1/2\,O_2 = FeO$	264	64.7	25–1377
$3\,Fe + 2\,O_2 = Fe_3O_4$	111	321	25 – 600
	1093	299	600 – 1537
$2\,Fe + 3/2\,O_2 = Fe_2O_3$	818	257	25–680
	807	244	680–1537
$3\,Fe + C = Fe_3C$	– 10.4	– 10.2	840–1537
$H_2 + 1/2\,O_2 = H_2O_{(g)}$	246	54.9	25–1700
$H_2 + 1/2\,S_{2(g)} = H_2S$	90.3	49.4	25–1500
$Ni + 1/2\,O_2 = NiO$	245	98.5	25–1452
$Ni + 1/2\,S_{2(g)} = NiS_{(s)}$	187	72.0	400–580
$1/2\,S_{2(g)} + O_2 = SO_2$	362	73.1	25–1700

Ellingham diagrams

A graphical representation of the standard free energy often proves useful. Figure 2.2 shows the evolution of the free energy of reaction as a function of temperature for the formation of a number of oxides. For the sake of comparison, the stoichiometries of the equations are expressed relative to a single mole of oxygen. Thus, for the formation of the monovalent (M_2O), bivalent (MO) and trivalent oxides (M_2O_3):

$$4\,M + O_2 = 2\,M_2O \tag{2.16}$$

$$2\,M + O_2 = 2\,MO \tag{2.17}$$

$$4/3\,M + O_2 = 2/3\,M_2O_3 \tag{2.18}$$

Representations such as those given in Figure 2.2 are called Ellingham diagrams.

The position of the oxides in the Ellingham diagram permits a comparison of their relative thermodynamic stabilities. In Figure 2.2, the lower the position, the more stable the oxide. For example, the curve for the oxide Al_2O_3 is situated below that of Cr_2O_3 for all temperatures. Metallic aluminium is therefore capable of reducing the oxide Cr_2O_3 resulting in the formation of Al_2O_3 and Cr.

© 2007, first edition, EPFL Press

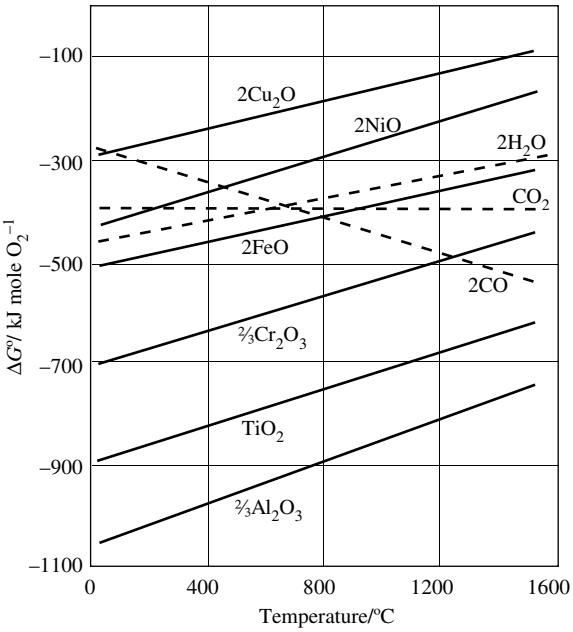


Figure 2.2 Ellingham diagram for a number of oxides.

Those metals with a negative free energy of oxide formation should react spontaneously with oxygen. However, in order to establish a state of equilibrium for the metal in the presence of its oxide and oxygen, contact must exist between the three phases and the interfacial reactions must be rapid. Fortunately, these conditions are not generally satisfied. At ambient temperatures, the oxidation normally produces a thin layer of oxide that physically separates the metal from the gaseous oxygen; the reaction stops if this barrier is not porous. However at high temperature, the increased ion mobility in the oxide layer allows the reaction to proceed further.

With the values shown in Figure 2.2, it is possible to calculate the partial pressure of oxygen in equilibrium with the oxide and the metal. Convention (I) applied to an ideal gas allows us to replace a_{O_2} by P_{O_2}, and Convention (II) allows us to set the activities of the metal, a_M and of the oxide, a_{oxide} to 1, because these are assumed to be pure.

$$\Delta G^\circ = -RT \ \ln K = RT \ \ln P_{O2} \tag{2.19}$$

Therefore, for each oxide at a given temperature, there is a corresponding oxygen partial pressure, called the ***dissociation pressure***.

The Ellingham diagram of oxides is also useful for the calculation of partial pressures of hydrogen and water vapor at equilibrium. At 600°C, can water vapor corrode copper?

$$2\ Cu + H_2O = Cu_2O + H_2 \tag{2.20}$$

© 2007, first edition, EPFL Press

If water vapor and hydrogen are the only gases present, the sum of their partial pressures is equal to the total pressure: $P = P_{H2} + P_{H2O}$. For a total pressure of 1 bar, the equilibrium constant of reaction (2.20) is expressed as:

$$K = \frac{P_{H_2}}{P_{H_2O}} = \frac{P_{H_2}}{1 - P_{H_2}} \tag{2.21}$$

We then find in Figure 2.2:

$$4\,Cu + O_2 = Cu_2O \qquad \Delta G^\circ \approx -230 \text{ kJ/mole} \tag{2.22}$$

$$2\,H_2 + O_2 = 2\,H_2O \qquad \Delta G^\circ \approx -400 \text{ kJ/mole} \tag{2.23}$$

In order to obtain reaction (2.20), (2.23) can be subtracted from (2.22), and the result divided by 2. In an analogous fashion, (2.20) yields $\Delta G^\circ = 85$ kJ mol^{-1}. With (2.8), the equilibrium constant (2.21) is found to be $K = 8 \times 10^{-6}$. This low value for K indicates that, at equilibrium, the partial pressure of hydrogen is extremely low: $P_{H2} = 10^{-6}$ bar. At 600 °C therefore, metallic copper should not be oxidized by contact with water vapor.

2.1.3 Activity of electrolytes

In the case of aqueous corrosion, one must often consider dissolved ***ionic species***. Unfortunately, their activities are rarely known. Indeed, because cations cannot be separated from the anions, the activity of an individual ion is not measurable. By defining the ***mean ionic activity,*** a measurable entity, one can get around this problem.

Thus, for a ***binary electrolyte***, that is, an electrolyte that contains only a single species of cation and anion, one obtains

$$a_\pm = c_\pm \cdot f_\pm \tag{2.24}$$

where $c_\pm$ corresponds to the mean ionic concentration. For the salt $A_{\nu+}B_{\nu-}$ where ν_+ and ν_- designate the stoichiometric coefficients of the cation and of the anion, respectively, the mean ionic concentration is expressed as:

$$c_\pm = (c_+^{\nu+} \cdot c_-^{\nu-})^{(1/(\nu_+ + \nu_-))} \tag{2.25}$$

In an analogous manner, $f_\pm$ represents the mean ionic activity coefficient:

$$f_\pm = (f_+^{\nu+} \cdot f_-^{\nu-})^{(1/(\nu_+ + \nu_-))} \tag{2.26}$$

The absorption equilibrium of hydrochloric acid in water illustrates this concept:

$$HCl_{(g)} = H^+_{(aq)} + Cl^-_{(aq)} \tag{2.27}$$

With (2.8), we obtain:

© 2007, first edition, EPFL Press

$$\Delta G^\circ = -RT \ \ln \frac{f_+ \cdot c_{H^+} \cdot f_- \cdot c_{Cl^-}}{P_{HCl}} = -RT \ \ln \frac{f_\pm^2 \cdot c_{HCl}^2}{P_{HCl}} \tag{2.28}$$

The activity of HCl in solution is thus proportional to the square root of the partial pressure:

$$a_\pm = f_\pm \cdot c_{HCl} \propto P_{HCl}^{1/2} \tag{2.29}$$

The mean activity coefficient of hydrochloric acid, $f_\pm$ is determined by measuring the concentration of HCl in solution as a function of the partial pressure p_{HCl}.

The definition of certain quantities that play an important role in corrosion and in electrochemistry, notably the equilibrium potential of an electrode (Section 2.2) and the pH, require knowledge about the activity of individual ions. By invoking certain non–thermodynamic hypotheses, it is possible to estimate these in certain cases. We present here two examples, namely the establishment of the pH scale and the estimation of activity coefficients in weakly concentrated aqueous solutions.

The pH is defined by equation (2.30):

$$\text{pH} = -\log a_{H^+} \tag{2.30}$$

Because the activity of the proton is not measurable in an absolute manner, buffer solutions are used as reference points to the ***pH scale***; to each buffer solution a pH value is assigned by convention (Table 2.3).

Table 2.3 pH values at 25 °C of several buffer solutions used as standards (adapted from ref. [3]).

Solution	pH
Potassium tetra–oxalate 0.05M	1.68
Saturated KH–tartrate	3.56
KH phthalate 0.05 M	4.01
KH_2PO_4 0.025M + Na_2HPO_4 0.025M	6.86
Borax 0.01M	9.18

For weakly concentrated solutions, (< 0.01 mol l^{-1}), it is possible to calculate values for the individual ionic activity coefficients by a microscopic approach using the ***Debye–Hückel theory***. This theory treats electrostatic interactions between ions as if each ion were surrounded by a "cloud" of ions of opposite charge [4]. The result of the calculation (2.32) shows that an individual ionic activity coefficient depends on the ***ionic strength*** of the solution J, expressed in mol l^{-1} and defined by (2.31).

$$J = \frac{1}{2} \sum z_i^2 \ c_i \tag{2.31}$$

$$\log f_i = -A_{DB} \cdot z_i^2 \cdot J^{1/2} \tag{2.32}$$

© 2007, first edition, EPFL Press

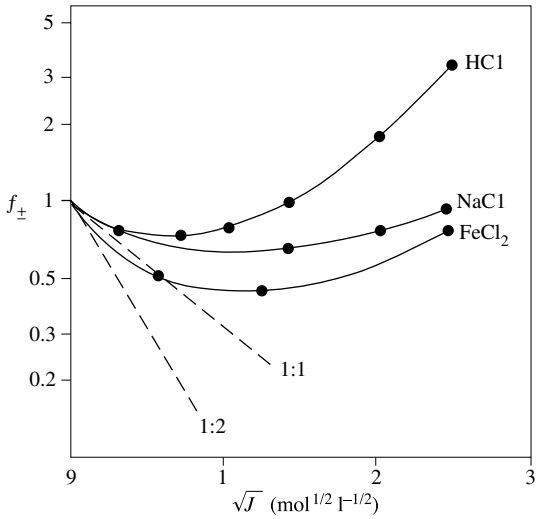


Figure 2.4 Mean ionic activity coefficient as a function of the ionic strength. The dashed lines indicate the Debye-Hückel law.

Here z_i represents the charge of the ion i, A_{DB} = 0.51 mol$^{-1/2}$l$^{1/2}$ at 25 °C and A_{DB} = 0.58 mol$^{-1/2}$ l$^{1/2}$ at 80 °C. With (2.26), the mean ionic activity coefficient of a binary electrolyte is obtained:

$$\log f_\pm = -A_{DB} \left| z_+ z_- \right| J^{1/2} \tag{2.33}$$

Figure 2.4 shows the value of the mean ionic activity coefficient for several binary electrolytes as a function of the square root of the ionic strength. The Debye–Hückel law represents the tangent of the curves at zero concentration. The equation (2.34), due to Güntelberg [4], gives a better estimation of the value of $f_\pm$ for the range of $0.001 < J < 0.01$ mol/l.

$$\log f_\pm = -\frac{A_{DB} \left| z_+ z_- \right| J^{1/2}}{1 + J^{1/2}} \tag{2.34}$$

In more heavily concentrated solutions, the activity coefficient increases with the ionic strength and may even exceed the value of $f_\pm$ = 1. Under these conditions, a theoretical prediction of its value is no longer possible.

2.2 ELECTROCHEMICAL EQUILIBRIUM

2.2.1 Standard potential of a cell reaction

In an electrochemical cell such as a battery or a fuel cell, the electrons that are freed at the anode pass via the external conductor to the cathode, where they participate

© 2007, first edition, EPFL Press

in the partial reduction reaction. A battery is thus capable of supplying electrical work. Equation (2.35) gives the electrical work, W_e, of a reversible reaction:

$$W_e = -n \cdot F \cdot E_{rev} \tag{2.35}$$

F is the Faraday constant (96,485 C/mol), n the charge number, and E_{rev} the ***reversible potential or equilibrium potential of the cell reaction***. By convention, the work supplied by a system is negative, which explains the negative sign of equation (2.35). For a given electrochemical redox reaction (2.1), the free energy of reaction is equal to:

$$\Delta G = \sum \nu_i \, \mu_i = -n \, F \, E_{rev} \tag{2.36}$$

which allows us to calculate the reversible electrical work (the maximum electrical work) that a battery or fuel cell can provide. In an analogous fashion, under standard conditions, we have

$$\Delta G^\circ = \sum \nu_i \, \mu_i^\circ = -n \, F \, E^\circ \tag{2.37}$$

where E° represents the standard potential of the cell reaction.

Formalism for representing electrochemical cells

For the representation of electrochemical cells, we use a set of symbols that distinguish between the different phases involved. The electrochemical cell of Figure 2.5 corresponds to that of Figure 1.2 under conditions of zero current. In the cell reaction (2.38), equivalent to (1.11), 4 moles of electrons are transferred per mole of oxygen (n = 4).

$$O_2 + 2\,H_2 = 2\,H_2O \tag{2.38}$$

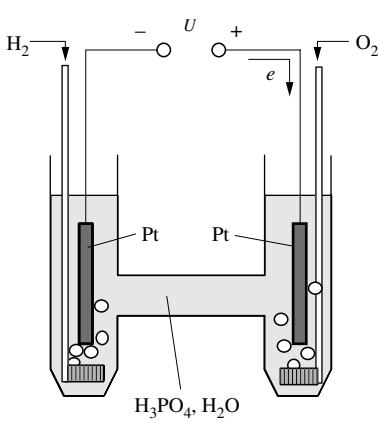


Figure 2.5 The H_2 / O_2 electrochemical cell.

© 2007, first edition, EPFL Press

This cell can be represented in the following way:

$$M' \mid \mathrm{Pt} \mid \mathrm{H_2}, \mathrm{H_3PO_4}, \mathrm{H_2O}, \mathrm{O_2} \mid \mathrm{Pt} \mid M'' \tag{2.39}$$

The vertical lines indicate the metal–metal and metal–electrolyte interfaces; M' and M'' represent the metal conductors (e.g., copper wires) connected to the measuring instrument (amperometer or voltmeter). The ***cell voltage***, U, is equal to the difference in electric potential between the metal of the right pole (M'') and that of the left pole (M'). Its value is therefore positive if the anode is located on the left.

$$U = \Phi_{M''} - \Phi_{M'} \tag{2.40}$$

Under standard conditions, the voltage at the terminals of an electrochemical cell at equilibrium is equal to the ***standard potential*** of the corresponding cell reaction. For example, for the reaction (2.38) taking place in the electrochemical cell (2.39):

$$U = E^\circ = -\Delta G^\circ / 4\,F = 1.23\,V \tag{2.41}$$

In reality, this equation is only valid in the absence of junction potentials (Section 2.4). In addition, it is often difficult, even impossible, to experimentally establish equilibrium conditions because corrosion reactions can occur, or because the kinetics of electrode reactions may be slow.

2.2.2 Standard potential of an electrode

Electrode reaction

An electric conductor, such as copper or platinum, in contact with an electrolyte is called an electrode. In electrochemistry, however, the term electrode is sometimes also used to designate an ***electrode reaction***, i.e. a charge–transfer reaction (oxidation or reduction) taking place at an electrode surface. For example, the terms "hydrogen electrode" (2.42) or "copper electrode" (2.43) designate the electrode reactions:

$$2\,\mathrm{H^+} + 2\,e = \mathrm{H_2} \tag{2.42}$$

$$\mathrm{Cu^{2+}} + 2\,e = \mathrm{Cu} \tag{2.43}$$

In the stoichiometric equation of an electrode reaction, the electrons are written explicitly. These reactions are also called ***half–cell reactions***. An electrochemical cell includes two half–cell reactions. By convention, electrode reactions are written as a reduction; this means that the electrons appear to the left in the equation:

$$\sum \nu_{\mathrm{ox,i}}\, B_{\mathrm{ox,i}} + n\,e = \sum \nu_{\mathrm{red,i}}\, B_{\mathrm{red,i}} \tag{2.44}$$

In this equation, $\nu_{\mathrm{ox,i}}$ represents the stoichiometric coefficients of the "oxidized" species $B_{\mathrm{ox,i}}$, and $\nu_{\mathrm{red,i}}$ indicates the stoichiometric coefficients of the "reduced" species $B_{\mathrm{red,i}}$. According to this definition, $\nu_{\mathrm{ox,i}}$ and $\nu_{\mathrm{red,i}}$ are always positive.

© 2007, first edition, EPFL Press

Equilibrium potential of an electrode

The ***equilibrium potential*** or ***reversible potential*** of an electrode is defined as the difference in electric potential between the metal and the solution at equilibrium. The electrode potential cannot be measured in an absolute manner; one can only measure *potential differences* between two electrodes in an electrochemical cell as schematically indicated in Figure 2.6.

This figure shows the levels of the electric potential in the electrochemical cell represented in Figure 2.5. The voltage measured at the contacts is given by $U = \Phi_{M''} - \Phi_{M'}$, where $\Phi_{M'}$ is the electric potential of the metal M' and $\Phi_{M''}$ that of the metal M''. This measurement in no case offers information concerning the differences of potential at the electrode–electrolyte interface, $\Phi_{M''} - \Phi_\sigma$ and $\Phi_{M'} - \Phi_\sigma$, because Φ_σ, the electric potential of the electrolyte, is not available.

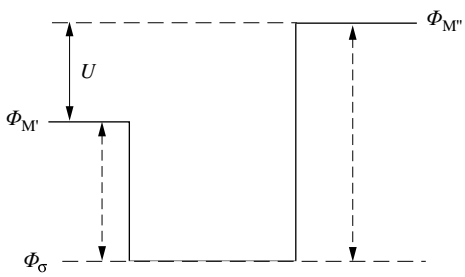


Figure 2.6 Levels of electric potential in an electrochemical cell without a liquid junction.

Standard hydrogen electrode

For the study of electrochemistry and corrosion one must be able to compare the equilibrium potentials of different electrode reactions. To these ends, *by convention,* a scale of standard electrode potentials is defined by arbitrarily assigning the value of zero to the equilibrium potential of the electrode (2.42), under standard conditions (P_{H2} = 1 bar =1.013 atm , T = 298 K, a_{H+} = 1):

$$2\,H^+(a_{H+}=1) + 2\,e = H_2 \qquad (P_{H2}=1\text{bar}) \qquad E^\circ = 0 \tag{2.45}$$

This reference is called the ***standard hydrogen electrode*** (she) or the normal hydrogen electrode (nhe).

By convention (2.45), the standard potential of an electrode reaction (2.44) becomes equal to the standard potential of the cell reaction (2.46).

$$\sum \nu_{ox,i}\, B_{ox,i} + \frac{n}{2} H_2 = \sum \nu_{red,i}\, B_{red,i} + n\,H^+ \tag{2.46}$$

For example, the standard potential of the copper electrode (2.43) corresponds to that of the cell reaction (2.47).

$$Cu^{2+} + H_2 = Cu + 2\,H^+ \qquad E^\circ = 0.340\ \text{V} \tag{2.47}$$

© 2007, first edition, EPFL Press

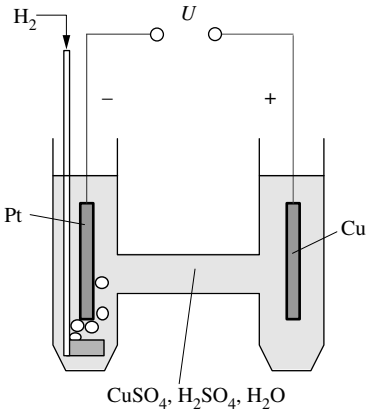


Figure 2.7 The H_2 / Cu electrochemical cell.

Figure 2.7 shows an electrochemical cell that corresponds to this equation. We have arbitrarily taken sulfate to be the anion.

A list of standard potentials of electrode reactions, organized by numerical value (Table 2.8), is referred to as *series of standard potentials* of electrode reactions or *electromotive force series.* A more complete set of such data can be found in reference [6].

When an electrode reaction is not found in the tables, its standard potential can be calculated by addition or subtraction of known reactions. If, for example, the standard free energy of the electrode reaction ΔG_3° corresponds to the sum of ΔG_1° and ΔG_2°,

$$\Delta G_3^\circ = \Delta G_1^\circ + \Delta G_2^\circ \tag{2.48}$$

it follows from (2.37) that the standard potential is equal to:

$$E_3^\circ = \frac{n(1)}{n(3)} E_1^\circ + \frac{n(2)}{n(3)} E_2^\circ \tag{2.49}$$

2.2.3 Variation of the standard potential as a function of temperature and pressure

Based on the relation (2.50), the variation of the standard potential of an oxidation–reduction reaction with temperature and pressure is given by (2.51) and (2.52), respectively:

$$\mathrm{d}\Delta G^\circ = \Delta V^\circ \mathrm{d}P - \Delta S^\circ \mathrm{d}T \tag{2.50}$$

$$\left[\frac{\partial E^\circ}{\partial T}\right]_P = -\frac{1}{nF}\left[\frac{\partial \Delta G^\circ}{\partial T}\right]_P = \frac{\Delta S^\circ}{nF} \tag{2.51}$$

© 2007, first edition, EPFL Press

Table 2.8 Standard potentials of electrode reactions.

Electrode	$E°$/V
$Li^+ + e = Li$	– 3.045
$Ca^{2+} + 2e = Ca$	– 2.84
$Na^+ + e = Na$	– 2.714
$Mg^{2+} + 2e = Mg$	– 2.356
$Al^{3+} + 3e = Al$	– 1.67
$Ti^{2+} + 2e = Ti$	– 1.63
$Mn^{2+} + 2e = Mn$	– 1.18
$Cr^{2+} + 2e = Cr$	– 0.90
$Cr^{3+} + 3e = Cr$	– 0.74
$Zn^{2+} + 2e = Zn$	– 0.76
$Fe^{2+} + 2e = Fe$	– 0.44
$Cd^{2+} + 2e = Cd$	– 0.403
$Ni^{2+} + 2e = Ni$	– 0.257
$Mo^{3+} + 3e = Mo$	– 0.20
$Sn^{2+} + 2e = Sn$	– 0.136
$Pb^{2+} + 2e = Pb$	– 0.125
$2\,H^+ + 2e = H_2$	0
$H_2MoO_4 + 6\,H^+ + 6\,e = Mo + 4H_2O$	0.114
$Cu^{2+} + 2e = Cu$	0.340
$H_2CrO_4 + 6\,H^+ + 6\,e = Cr + 4\,H_2O$	0.326
$O_2 + 2H_2O + 4e = 4OH^-$	0.401
$Cu^+ + e = Cu$	0.520
$Hg_2^{2+} + 2e = 2Hg$	0.796
$Ag^+ + e = Ag$	0.799
$Pd^{2+} + 2e = Pd$	0.915
$Pt^{2+} + 2e = Pt$	1.188
$O_2 + 4H^+ + 4e = 2H_2O$	1.229
$Au^{3+} + 3e = Au$	1.52
$Au^+ + e = Au$	1.83

© 2007, first edition, EPFL Press

$$\left[\frac{\partial E^\circ}{\partial P}\right]_T = -\frac{1}{nF}\left[\frac{\partial \Delta G^\circ}{\partial P}\right]_T = \frac{\Delta V^\circ}{nF} \tag{2.52}$$

In these equations, ΔS° represents the standard entropy of reaction and ΔV° the volume change during reaction. ΔS° is obtained from (2.53), where S_i° is the entropy of formation of the individual neutral and ionic species that participate in the reaction (the standard entropy of a perfect solid at 0 K is defined as having the value of zero) and V_i^m represents their molar volumes:

$$\Delta S^\circ = \sum \nu_i \, S_i^\circ \tag{2.53}$$

$$\Delta V^\circ = \sum \nu_i \, V_i^m \tag{2.54}$$

Because thermodynamic data for ionic species are not directly available, one arbitrarily sets the values for the solvated proton to zero,

$$\Delta G^\circ_{H^+} = \Delta H^\circ_{H^+} = S^\circ_{H^+} = 0 \tag{2.55}$$

which in turn allows the parameters for the other ionic species to be determined (Table 2.9).

Table 2.9 Thermodynamic properties of ionic and neutral species at 25 °C (adapted from ref. [6]).

Species	ΔH° (kJ mol^{-1})	ΔG° (kJ mol^{-1})	S° (J mol^{-1} K^{-1})
$H^+_{(aq)}$	0	0	0
$H_{2(g)}$	0	0	130.7
$O_{2(g)}$	0	0	205.0
$O_{2(aq)}$	–11.7	16.3	110.9
$H_2O_{(l)}$	–285.8	–237.2	69.9
$OH^-_{(aq)}$	–230.0	–157.3	–10.75
$Cu_{(s)}$	0	0	33.2
$Cu^+_{(aq)}$	72.1	50.3	41
$Cu^{2+}_{(aq)}$	65.8	65.7	–97.2
$Ag_{(s)}$	0	0	42.6
$Ag^+_{(aq)}$	105.6	77.2	72.7
$AgCl_{(s)}$	–127.1	–109.8	96.3
Cl^-	–166.9	–131.1	56.7

© 2007, first edition, EPFL Press

The standard potential of cell reactions varies only slightly with temperature (in general ≤ 1 mV/°C). For example, for the reaction (2.47) it is found:

$$\Delta S^\circ = S^\circ_{\mathrm{Cu(s)}} + 2\,S^\circ_{\mathrm{H^+}} - S^\circ_{\mathrm{Cu^{2+}}} - S^\circ_{\mathrm{H_2}}$$

$$= 33.2 + 97.2 - 130.7 = -0.3 \ \ \mathrm{J\,mol^{-1}\,K^{-1}} \tag{2.56}$$

With (2.51) we obtain thus for the temperature coefficient of the standard potential:

$$\left(\partial E^\circ / \partial T\right)_{\mathrm{P}} = -1.5 \times 10^{-6} \ \mathrm{V/°C}$$

Integration of this relationship allows us to calculate E° as a function of temperature. For the reaction (2.47), the result of the integration corresponds to the standard potential of the copper electrode, measured with respect to that of the hydrogen electrode at the same temperature.

The influence of pressure on the standard potentials of cell reactions is negligible, except for those reactions that involve a gas. For example, for reaction (2.47) one finds:

$$\Delta V^\circ = V^{\mathrm{m}}_{\mathrm{Cu_{(s)}}} + 2\,V^{\mathrm{m}}_{\mathrm{H^+}} - V^{\mathrm{m}}_{\mathrm{Cu^{2+}}} - V^{\mathrm{m}}_{\mathrm{H_2}} \cong -V^{\mathrm{m}}_{\mathrm{H_2}} \tag{2.57}$$

With the assumption that hydrogen behaves as a perfect gas, we obtain:

$$\left(\frac{\partial E^\circ}{\partial P}\right)_{\mathrm{T}} = \frac{R\,T}{n\,F\,P_{\mathrm{H_2}}}$$

This relation describes the variation with pressure of the standard potential of an electrode reaction measured against a standard hydrogen electrode at the same pressure.

2.2.4 Nernst equation

The Nernst equation applies to electrodes at equilibrium. It permits the calculation of the equilibrium potential as a function of concentration (activity) and temperature. The starting point for deriving the Nernst equation of an electrode reaction is the cell reaction (2.46),

$$\sum \nu_{\mathrm{ox,i}}\,B_{\mathrm{ox,i}} + \frac{n}{2}\mathrm{H_2} = \sum \nu_{\mathrm{red,i}}\,B_{\mathrm{red,i}} + n\,\mathrm{H^+} \tag{2.46}$$

whose free energy of reaction is equal to:

$$\Delta G = \Delta G^\circ + RT\,\ln \frac{a^{\mathrm{n}}_{\mathrm{H^+}} \prod a^{\nu_{\mathrm{red,i}}}_{\mathrm{red,i}}}{P^{\mathrm{n/2}}_{\mathrm{H_2}} \prod a^{\nu_{\mathrm{ox,i}}}_{\mathrm{ox,i}}} \tag{2.58}$$

© 2007, first edition, EPFL Press

With $\Delta G = -\,n\,F\,E_{rev}$, we obtain the reversible potential of the redox reaction (2.46):

$$E_{rev} = E^\circ - \frac{RT}{n\,F}\,\ln\frac{a_{H^+}^{n}\prod a_{red,i}^{\nu_{red,i}}}{P_{H_2}^{n/2}\prod a_{ox,i}^{\nu_{ox,i}}} = E^\circ + \frac{RT}{n\,F}\,\ln\frac{P_{H_2}^{n/2}\prod a_{ox,i}^{\nu_{ox,i}}}{n\,F\,a_{H^+}^{n}\prod a_{red,i}^{\nu_{red,i}}} \tag{2.59}$$

When $a_{H+} = 1$ and $P_{H2} = 1$ atm, the reversible potential of (2.46) becomes:

$$E_{rev} = E^\circ + \frac{RT}{n\,F}\,\ln\frac{\prod a_{ox,i}^{\nu_{ox,i}}}{\prod a_{red,i}^{\nu_{red,i}}} \tag{2.60}$$

This potential is equal to the reversible potential of the electrode reaction (2.44).

$$\sum \nu_{ox,i}\,B_{ox,i} + n\,e = \sum \nu_{red,i}\,B_{red,i} \tag{2.44}$$

Equation (2.60) represents the general form of the ***Nernst equation of an electrode reaction***. In this equation, $a_{ox,i}$ and $a_{red,i}$ express the activities of the "oxidized" species (those species that appear on the side of the equality with the electrons) and the "reduced" species (appearing on the side without the electrons) in (2.44).

For the electrode reaction

$$Cu^{2+} + 2\,e = Cu \tag{2.43}$$

the Nernst equation (2.60) reads:

$$E_{rev} = 0.340 + \frac{RT}{2\,F}\,\ln a_{Cu^{2+}} \tag{2.61}$$

At $T = 25$ °C and after conversion to base–10 logarithms, we obtain:

$$E_{rev} = 0.340 + \frac{0.059}{2}\,\log a_{Cu^{2+}} \tag{2.62}$$

Equations (2.61) and (2.62) depend upon the activity of the Cu^{2+} ions, which is not measurable. In practice, this problem is most often avoided by a simplification: the activity is replaced by the concentration. Equation (2.62) then becomes:

$$E_{rev} \approx 0.340 + \frac{0.059}{2}\,\log c_{Cu^{2+}} \tag{2.63}$$

This assumption is valid as long as the concentration of Cu^{2+} ions is weak. However, even for higher concentrations, the error in the potential resulting from this simplification remains relatively small, in general, because the activity appears only in the logarithmic term.

A comment about the sign

In the Nernst equation (2.60), the subscripts *ox* and *red* ("oxidized species" and "reduced species", equation (2.44)) replace the subscripts "reactant" and "product"

© 2007, first edition, EPFL Press

that we typically find in chemical thermodynamics. The advantage of this convention is that the sign of the half–cell reversible potential is independent of the manner in which the stoichiometric equation is written. In the case of the copper electrode, for example, the ion Cu^{2+} represents the oxidized species and the metallic Cu is the reduced species. With (2.60) the same value for E_{rev} is obtained by writing

$$Cu^{2+} + 2e = Cu \qquad \text{(reduction)}$$

or

$$Cu = Cu^{2+} + 2e \qquad \text{(oxidation)}$$

On the other hand, from a strict thermodynamic point of view, electrode reactions should always be expressed as a reduction in order to yield the general form of the Nernst equation (2.58), as described in Section 2.2.2.

2.3 EQUILIBRIUM POTENTIAL AND CORROSION BEHAVIOR

2.3.1 Galvanic effects

Galvanic interactions between different metals may accelerate the rate of corrosion of the less noble metal (Chapter 7). The Nernst equation offers a means of estimating the importance of these effects according to thermodynamic criteria. Two distinct cases are considered here:

- the formation of an electrochemical cell between two metals, each of which is in contact with a solution of its own ions;
- the displacement reaction between a metal and a solution containing the cations of another metal.

Theoretical voltage of an electrochemical cell

The theoretical voltage of an electrochemical cell is equal to the difference between the reversible potentials of the two electrode reactions, as calculated by the Nernst equation. For the *Daniell cell* (Figure 2.10), the theoretical voltage is given by:

$$\Delta E = E_{rev,Cu} - E_{rev,Zn} \tag{2.64}$$

where

$$E_{rev,Cu} = 0.340 + \frac{RT}{2F} \ln a_{Cu^{2+}} \tag{2.65}$$

and

$$E_{rev,Zn} = -0.760 + \frac{RT}{2F} \ln a_{Zn^{2+}} \tag{2.66}$$

© 2007, first edition, EPFL Press

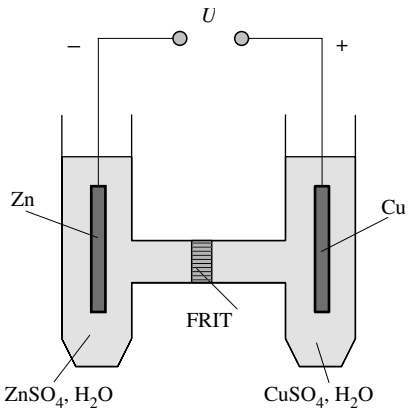


Figure 2.10 The Daniell Cell

In this case, the theoretical voltage ΔE differs from the measured cell voltage U, because a junction potential forms at the interface between the zinc sulfate and copper sulfate solutions (Section 2.4).

Displacement reactions

If a less noble metal comes into contact with a solution containing ions of a more noble one, a displacement reaction takes place: the less noble metal dissolves, while a deposit of the more noble metal forms on the surface. In this fashion, a bar of zinc lowered into a solution of copper sulfate is coated spontaneously by a copper deposit (Figure 2.11):

$$Zn + Cu^{2+} \rightarrow Cu + Zn^{2+} \tag{2.67}$$

With the Nernst equation, the concentrations of the ions at equilibrium can be calculated. Because the deposit that forms is usually porous, the system behaves like

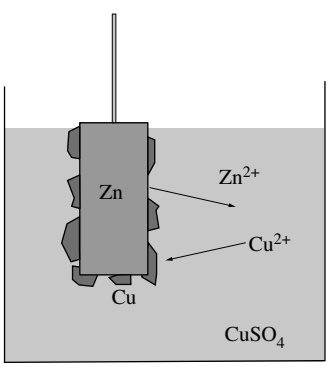


Figure 2.11 Displacement reaction between copper and zinc.

© 2007, first edition, EPFL Press

a short–circuited electrochemical cell, consisting of a copper electrode in contact with a solution of $CuSO_4$, and of a zinc electrode in contact with a solution of $ZnSO_4$. Equilibrium is reached as soon as

$$E_{\text{rev,Cu}} = E_{\text{rev, Zn}} \tag{2.68}$$

With the Nernst equation, the ratio of the activities at equilibrium can be determined.

$$\frac{a_{\text{Cu}^{2+}}}{a_{\text{Zn}^{2+}}} = \exp\left(-\frac{2F}{RT}\left(E^{\circ}_{\text{Cu}} - E^{\circ}_{\text{Zn}}\right)\right) \tag{2.69}$$

At T = 25 °C, we find using the values of Table 2.8: $a_{\text{Cu2+}}/a_{\text{Zn2+}} = \exp(-86) \approx 0$. Leaving aside the distinction between activity and concentration, we can state that at equilibrium, the concentration of copper ions in solution is practically zero.

The displacement reaction of copper with iron is used to recover copper ions in waste water. Displacement reactions may also cause corrosion. A well–known example concerns heating systems; copper ions liberated by corrosion of a hot–water heater made of copper react downstream with the wall of a zinc–coated steel pipe. The microscopic deposits of metallic copper form a galvanic cell with the wall and thus accelerate locally the rate of corrosion.

2.3.2 Protection potential

The Nernst equation describes the equilibrium condition of an electrode. For example, the equlibrium potential of a copper electrode in contact with a solution of Cu^{2+} ions at a concentration (activity) of 10^{-2} mol l^{-1} is equal to

$$E_{\text{rev}} = 0.34 + \frac{0.059}{2}\log a_{\text{Cu}^{2+}} = 0.28\,\text{V} \tag{2.70}$$

The copper will not dissolve unless the potential of the electrode exceeds this value.

In practice, a metal being part of a structure or a machine is rarely in contact with a solution of its own ions, at a uniform and known concentration. More often, metals are in contact with an electrolyte that does not contain their own ions. In this case, the reversible potential remains undefined. For example, according to (2.65) the reversible potential of a copper electrode in contact with sulfuric acid is equal to:

$$E_{\text{rev}} = -\infty \tag{2.71}$$

From a thermodynamics point of view, the copper should therefore corrode whatever its potential. However, as soon as the metal corrosion begins, Cu^{2+} ions enter the solution at the electrode surface and the concentration of the ions instantaneously exceeds zero. The reversible potential, as shown in (2.65), takes on a finite value that increases as the Cu^{2+} concentration at the electrode surface increases. This concentration depends on the rate of transport of the Cu^{2+} ions towards the bulk solution. Figure 2.12 schematically illustrates this behavior (which does not in any case refer to a true equilibrium situation because corrosion occurs).

© 2007, first edition, EPFL Press

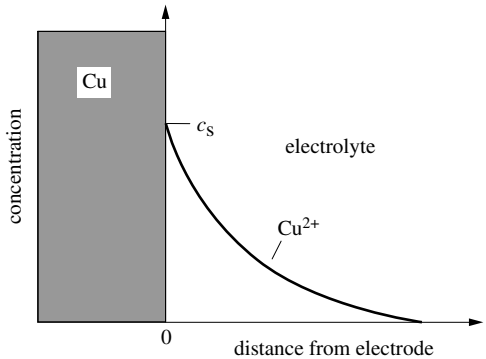


Figure 2.12 Copper ion concentration profile near the electrode during copper dissolution.

The rate of corrosion v_{cor} expressed in mol m^{-2} s^{-1}, is proportional to the Cu^{2+} ion concentration at the surface, c_s.

$$v_{cor} = k_L\, c_s \tag{2.72}$$

Here, k_L designates the mass transfer coefficient. Its value depends on the diffusion coefficient and on convection (Section 4.3). In aqueous solutions, k_L typically has a value on the order of 10^{-4} to 10^{-5} m s^{-1}. Introducing (2.72) into (2.65) and replacing the activity of Cu^{2+} with the concentration, one finds a relation between the reversible potential of an electrode and its rate of corrosion:

$$E_{rev} = 0.34 + \frac{0.059}{2} \log \frac{v_{cor}}{k_L} \tag{2.73}$$

In this case, v_{cor} represents the *maximum rate* of corrosion, because we have made the assumption that the metal surface is in equilibrium with the copper ions in its immediate vicinity, thus implying rapid charge transfer at the interface.

A corrosion rate of $v_{cor} = 10^{-7}$ mole m^{-2} s^{-1} is a very low rate. For copper, this corresponds to a loss of thickness due to corrosion of about 22 µm per year. For many purposes (but not always!) this is negligible. Assuming a mass transfer coefficient of $k_L = 10^{-4}$ m s^{-1}, (2.73) yields a reversible potential of $E_{rev} = 0.163$ V. With (2.72), this corresponds to a concentration of Cu^{2+} ions at the electrode surface of $c_s = 10^{-6}$ mol l^{-1}. By turning around the reasoning, it follows that the rate of corrosion for copper is very slow if the potential of the metal is equal or inferior to 0.163 V. At equilibrium, this potential corresponds to a copper ion concentration of 10^{-6} mol l^{-1}.

In general, the maximum rate of corrosion does not exceed approximately 20 µm per year if the dissolved ion concentration at the metal surface remains equal or inferior to 10^{-6} mol l^{-1}. A rate of corrosion of this magnitude can often be tolerated. Thus, we define the ***protection potential*** (E_{prot}) as the reversible potential predicted by the Nernst equation for a dissolved metal–ion activity (concentration) of 10^{-6} mol l^{-1}.

© 2007, first edition, EPFL Press

$$E_{\text{prot}} = E^\circ + \frac{RT}{nF} \ln 10^{-6} \tag{2.74}$$

The protection potential is not a true thermodynamic quantity, in contrast to the reversible potential, because it does not refer to an equilibrium situation. The definition is somewhat arbitrary, based in part on non–thermodynamic considerations.

The protection potential plays an important role in engineering because it defines the electrochemical conditions that "protect" a metal against corrosion. If the potential of the metal is at or below the protection potential, the rate of corrosion can be considered negligible. If required, the definition of the protection potential can be adapted to the precise conditions of a given situation by using another value for the surface concentration that defines the maximum rate of corrosion that can be tolerated.

2.3.3 Equilibrium between ions of different oxidation state

Most metals may form ions of different oxidation states. The Nernst equation allows us to predict the relative stability of these different ions in contact with an electrode. Three distinct cases will be considered:

- disproportionation of thermodynamically unstable ions;
- corrosion by metal ions having more than one oxidation state;
- equilibrium between dissolved species in contact with an inert electrode.

Disproportionation reactions

Often several oxidation states of the same ion can coexist in solution, such as Cu^+ and Cu^{2+}, Fe^{2+} and Fe^{3+}, or Cr^{2+}, Cr^{3+} and CrO_4^{2-}. A ***disproportionation reaction*** is defined as an oxidation–reduction reaction between two identical ions, producing simultaneously species with higher and lower valences.

In a solution that contains no complexing agents, Cu^+ ions can disproportionate into Cu et Cu^{2+}:

$$2\,Cu^+ = Cu + Cu^{2+} \tag{2.75}$$

The overall reaction consists of two partial reactions:

$$Cu^+ + e = Cu \qquad E^\circ = 0.520\ \text{V} \tag{2.76}$$

$$Cu^{2+} + 2\,e = Cu \qquad E^\circ = 0.340\ \text{V} \tag{2.77}$$

At equilibrium, the reversible potentials of the two reactions are equal. With the Nernst equation, the activities of the Cu^+ and Cu^{2+} ions can be calculated:

$$E_{\text{rev}} = 0.520 + \frac{RT}{F} \ln a_{Cu^+} = 0.340 + \frac{RT}{2F} \ln a_{Cu^{2+}} \tag{2.78}$$

© 2007, first edition, EPFL Press

Using $T = 298$ K, we obtain:

$$\frac{a^2_{Cu^+}}{a_{Cu^{2+}}} = 7.95 \times 10^{-7} \text{ mol l}^{-1} \tag{2.79}$$

By replacing the activities with concentrations, we can calculate the fraction of Cu^+: $X_{Cu+} = c_{Cu+}/(c_{Cu+} + c_{Cu2+})$. Because metallic copper deposits out:

$$X_{Cu^+} + X_{Cu^{2+}} = 1 \tag{2.80}$$

Thus with (2.79):

$$c_{tot} \cdot \frac{X^2_{Cu^+}}{1 - X_{Cu^+}} = 7.95 \times 10^{-7} \text{ mol l}^{-1} \tag{2.81}$$

For $X_{Cu+} << 1$, this yields:

$$X_{Cu^+} = \frac{8.9 \times 10^{-4}}{\sqrt{c_{tot}}} \text{ mol l}^{-1} \tag{2.82}$$

This result shows that the fraction of Cu^+ ions at equilibrium is generally very small. When, in the absence of complexing agents, a salt of monovalent copper is dissolved in water, the Cu^+ ions therefore disproportionate according to reaction (2.75). In a general way, we may conclude that if a metal can form two ions of different oxidation state in solution, the lower valence ion is thermodynamically unstable if its equilibrium potential is more noble than that of the higher valence ion.

Corrosion by metal ions having more than one oxidation state

In aqueous solution, iron exists in two oxidation states: Fe^{2+} and Fe^{3+}. In contrast to copper, the standard potential for the reduction of the less oxidized species, Fe^{2+}, is the less noble of the two:

$$Fe^{2+} + 2e = Fe \qquad E° = -0.440 \text{ V} \tag{2.83}$$

$$Fe^{3+} + 3e = Fe \qquad E° = -0.037 \text{ V} \tag{2.84}$$

According to these values, only the Fe^{2+} ion is capable of coexisting in equilibrium with metallic iron, with the fraction of Fe^{3+} being negligible . In the absence of oxygen, the anodic dissolution of iron therefore produces Fe^{2+} rather than Fe^{3+}. On the other hand, ferric ions Fe^{3+}, in a solution of $FeCl_3$ for example, corrode metallic iron.

$$Fe + 2\,Fe^{3+} = 3\,Fe^{2+} \tag{2.85}$$

© 2007, first edition, EPFL Press

Fe^{3+} is a common oxidizing agent used for chemical etching of precision parts in microtechnology and for the accelerated simulation of corrosion in the laboratory.

Equilibrium between dissolved species in contact with an inert electrode

In the absence of metallic iron or of another non–noble metal, ferric and ferrous ions can coexist in solution. It is therefore possible to measure the reversible potential of the electrode reaction (2.86) using an inert platinum electrode. The value depends on the relative concentration of the Fe^{3+} and Fe^{2+} ions.

$$Fe^{3+} + e = Fe^{2+} \qquad E^\circ = 0.771\ \text{V} \tag{2.86}$$

$$E_{rev} = 0.771 + \frac{RT}{F} \ln \frac{a_{Fe^{3+}}}{a_{Fe^{2+}}} \tag{2.87}$$

Figure 2.13 shows schematically the difference in reaction when iron or platinum electrodes are employed. In principle, ferric ions can corrode any metal whose reversible potential is less than the potential of the electrode reaction (2.86).

The potential of a half–cell reaction between dissolved species is called the oxidation–reduction potential or ***redox potential***. It describes the oxidative strength of a solution. For example, the electrode reaction (2.88) sets the redox potential of an aqueous solution containing dissolved oxygen.

$$O_2 + 4\,H^+ + 4\,e = 2\,H_2O \qquad E^\circ = 1.23\ \text{V} \tag{2.88}$$

The very high standard potential indicates that oxygen is a powerful oxidizing agent. At $T = 298$ K and by setting $pH = -\log a_{H+}$, the reversible potential of reaction (2.88) becomes:

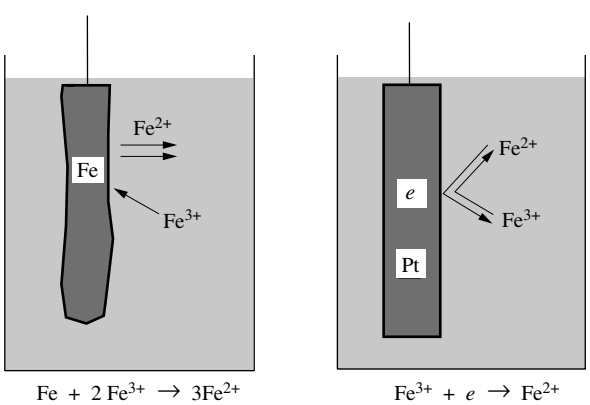


Figure 2.13 Reduction of Fe^{3+} ions on an iron electrode undergoing corrosion and on an inert electrode made of platinum.

© 2007, first edition, EPFL Press

$$E_{rev} = 1.23 - 0.059 \cdot pH + \frac{0.059}{4} \log P_{O_2} \tag{2.89}$$

According to (2.89), the oxidative strength of dissolved oxygen decreases with an increase in the pH. Table 2.14 gives the standard potentials of different electrode reactions between dissolved species. It shows that hydrogen peroxide, H_2O_2, and bichromate ions, $Cr_2O_7^{2-}$, are particularly strong oxidizing agents.

Table 2.14 Standard potentials for electrode reactions between dissolved species.

Reaction	E^o / V
$H_2O_2 + 2\ H^+ + 2\ e = 2\ H_2O$	1.763
$Cr_2O_7^{2-} + 14\ H^+ + 6\ e = 2\ Cr^{3+} + 7\ H_2O$	1.36
$Fe^{3+} + e = Fe^{2+}$	0.771
$Fe(CN)_6^{3-} + e = Fe(CN)_6^{4-}$	0.361
$TiO^{2+} + 2H^+ + e = Ti^{3+} + H_2O$	0.100
$TiO^{2+} + 2\ H^+ + 2e = Ti^{2+} + H_2O$	–0.135
$Cr^{3+} + e = Cr^{2+}$	–0.424

2.3.4 Corrosion in the presence of complexing agents

Complexing agents replace water molecules in the hydration sphere of a dissolved metal ion. Their presence changes the stoichiometry of the dissolution reaction of the metal and generally leads to a lower standard potential (Table 2.15) facilitating corrosion.

When a metal dissolves in the presence of complexing agents ions whose oxidation state corresponds to the most stable complex are predominantly formed. For example, the dissolution of copper in a non–complexing environment produces the Cu^{2+} ion, but in the presence of the complexing agents Cl^- or CN^-, the monovalent species $CuCl_2^-$ and $Cu(CN)_2^-$ are formed.

Table 2.15 Standard potentials of electrode reactions involving complexing agents.

Reaction	E^o/V
$Fe(CN)_6^{4-} + 2\ e = Fe + 6\ CN^-$	– 1.56
$Fe(CN)_6^{3-} + 3\ e = Fe + 6\ CN^-$	– 0.92
$Au(CN)_2^- + e = Au + 2\ CN^-$	– 0.595
$Cu(CN)_2^- + e = Cu + 2\ CN^-$	– 0.44
$Ag(CN)_2^- + e = Ag + 2\ CN^-$	– 0.31
$Cu(NH_3)_2^+ + e = Cu + 2\ NH_3$	– 0.10
$CuCl_2^- + e = Cu + 2\ Cl^-$	0.225
$AuCl_4^- + 3\ e = Au + 4\ Cl^-$	1.002

© 2007, first edition, EPFL Press

Gold: a particularly good example

In an aqueous environment devoid of complexing agents, gold is present in monovalent (Au^+) and trivalent (Au^{3+}) states. The standard equilibrium potentials favor the Au^{3+} ion.

$$Au^{3+} + 3e = Au \qquad E° = 1.52 \text{ V} \tag{2.90}$$

$$Au^{+} + e = Au \qquad E° = 1.83 \text{ V} \tag{2.91}$$

These high values of $E°$ confirm the fact that gold is a noble metal that does not easily corrode.

The presence of cyanide ions greatly alters this behavior because they form a very stable complex with monovalent gold. The electrode reaction between metallic and monovalent gold now reads:

$$Au(CN)_2^- + e = Au + 2\ CN^- \qquad E° = -0.595 \text{ V} \tag{2.92}$$

The low value of this standard potential indicates that in the presence of cyanide ions, gold no longer behaves as a noble metal. An oxidizing agent, such as dissolved oxygen, is capable of corroding it relatively easily.

These phenomena are exploited in the mining of gold. Gold is present in mineral deposits in metallic form, but at very dilute concentration. In order to extract it, the rock is ground and the resulting gold–containing slurry is treated with a cyanide solution in the presence of oxygen. The gold dissolves as the complex $Au(CN)_2^-$. Once a gold enriched solution is thus obtained, the complex is reduced to recover the gold metal.

Complexation constant

Starting with the standard potential of the electrode reaction $M^{n+} + n\ e = M$, it is possible, using the Nernst equation, to calculate the equilibrium constant for complex formation between the ion M^{n+} and a dissolved complexing agent. With the following calculation, the complexation constant for equation (2.93) is determined.

$$Au^{+} + 2\ CN^- = Au(CN)_2^- \tag{2.93}$$

$$K = \frac{a_{Au(CN)_2^-}}{a_{Au^+} \cdot a_{CN^-}^2} \tag{2.94}$$

$$a_{Au^+} = \frac{a_{Au(CN)_2^-}}{K\, a_{CN^-}^2} \tag{2.95}$$

© 2007, first edition, EPFL Press

The Nernst equation for the electrode reaction (2.91) yields therefore:

$$E_{rev} = E^\circ + \frac{RT}{nF} \ln \frac{a_{Au(CN)_2^-}}{K\, a^2_{CN^-}} = E^\circ_{complex} + \frac{RT}{nF} \ln \frac{a_{Au(CN)_2^-}}{a^2_{CN^-}} \tag{2.96}$$

$E^\circ_{complex}$ corresponds to the standard potential of reaction (2.92).

$$E^\circ_{complex} = E^\circ - \frac{RT}{nF} \ln K \tag{2.97}$$

In this example, $E^\circ_{complex}$ = –0.595 V, E° = 1.83 V and n = 1. Therefore, the complexation constant, at 25°C, has a value of: $K \approx 10^{41}$ mol^{-2} l^2. In the presence of cyanide ions, the concentration of non–complexed Au^+ ions is practically zero.

2.3.5 Pourbaix diagrams

The reversible potential of many electrode reactions, notably those involving oxides, depends on the pH. The ***potential – pH diagrams***, also called ***Pourbaix diagrams***, display the reversible potential, as calculated by the Nernst equation, as a function of pH [7].

For a given metal, the diagram is drawn by taking account of all possible electrode and chemical reactions at 25 °C:

- electrochemical equilibrium between a metal and its ions:

$$Mn^{n+} + n\,e = M \tag{2.98}$$

- electrochemical equilibrium between a metal and its oxide:

$$MO_{n/2} + n\,H^+ + n\,e = M + \frac{n}{2} H_2O \tag{2.99}$$

- electrochemical equilibrium between two oxides of different oxidation states:

$$MO_{n/2} + n\,H^+ + n\,e = MO_{(n-1)/2} + \frac{n}{2}\, H_2O \tag{2.100}$$

- chemical equilibrium in acidic solution between an oxide and dissolved ions:

$$MO_{n/2} + n\,H^+ = MO^{n+} + \frac{n}{2}\, H_2O \tag{2.101}$$

- chemical equilibrium in basic solution between an oxide and dissolved ions:

$$MO_{n/2} + 2\,OH^- = MO^{2-}_{(n/2+1)} + H_2O \tag{2.102}$$

We will consider the potential–pH diagram of aluminum as an example.

© 2007, first edition, EPFL Press

The potential–pH diagram of aluminum

Table 2.16 gives the energy of formation for different ionic and neutral oxidation products of aluminum at 25 °C.

Table 2.16 Standard free energy of formation of aluminum compounds (adapted from ref. [6]).

Chemical species	ΔG^o/ kJ mol^{-1}
$Al^{3+}_{(aq)}$	– 485
$Al(OH)_{4(aq)}^{-}$	–1297.8
$Al(OH)^{2+}_{(aq)}$	–694.1
$OH^{-}_{(aq)}$	–157.3
H_2O	–237.2
$Al(OH)_3$ amorphous	–1137.6
$Al(OH)_3$ Gibbsite	–1154.9
$Al_2O_3 \bullet H_2O$ Boehmite	–1825.4

These data are used to calculate the reversible potentials and the equilibrium constants for the half cell reactions listed in Table 2.17 and shown in Figure 2.18.

It is assumed here that the formula $Al(OH)_3$ corresponds to gibbsite (also called hydrargillite), and that no other oxides or hydroxides form. The reversible potential is determined by the thermodynamic properties of the reaction products and thus by

Table 2.17 Halfcell reactions represented in Figure 2.18.

No.	Reaction	Equation
1	$Al^{3+} + 3e = Al$	$E_{rev} = -1.676 + \frac{0.059}{3} \log a_{Al^{3+}}$ $a_{Al^{3+}} = 10^{-6} \rightarrow E = -1.794\ V$
2	$Al(OH)_3 + 3H^+ + 3e = Al + 3H_2O$	$E_{rev} = -1.563 - 0.059\ pH$
3	$Al(OH)_3 + 3H^+ = Al^{3+} + 3H_2O$	$pH = 2.44 - \frac{1}{3} \log a_{Al^{3+}}$ $a_{Al^{3+}} = 10^{-6} \rightarrow pH = 4.44$
4	$Al(OH)_3 + OH^- = Al(OH)_4^-$	$pH = 16.53 + \log a_{Al(OH)_4^-}$ $a_{Al(OH)_4^-} = 10^{-6} \rightarrow pH = 10.53$
5	$Al(OH)_4^- + 3e = Al + 4OH^-$	$E_{rev} = -1.20 + \frac{0.059}{3} \log a_{Al(OH)_4^-} - 0.079\ pH$ $a_{Al(OH)_4^-} = 10^{-6} \rightarrow E = -1.32 - 0.079\ pH$

© 2007, first edition, EPFL Press

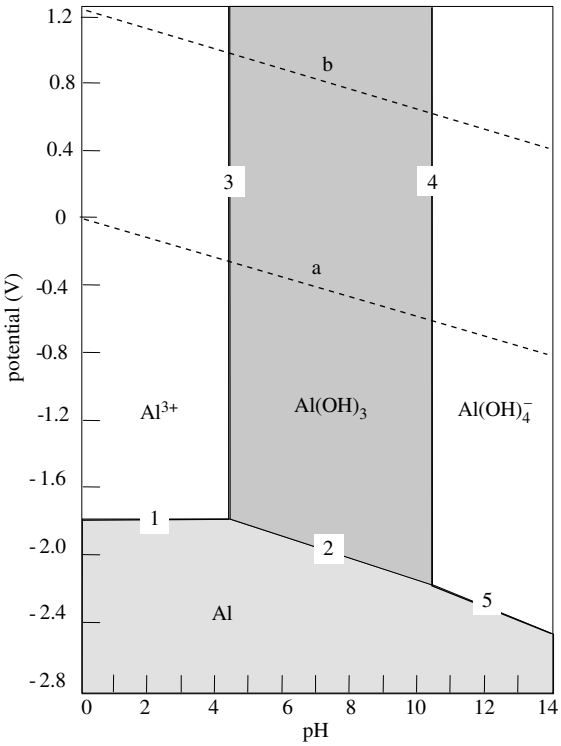


Figure 2.18 Potential-pH diagram for aluminum (T = 25 °C). The concentration of the dissolved ionic species, Al^{3+} and $Al(OH)_4^-$ are equal to 10^{-6} mol l^{-1}. The numbers refer to the reactions given in Table 2.17.

their chemical nature. Unfortunately, we do not always have this information for the oxide layers that form on metals.

In order to draw the potential–pH diagram for aluminum, it is necessary to define concentrations of the dissolved species (Al^{3+} and $Al(OH)_4^-$) at equilibrium. For these purposes, we will use here the value of $c = 10^{-6}$ mol/l, but similar diagrams could be drawn for any concentration. Figure 2.18 shows the resulting potential–pH diagram. Because the concentration of the dissolved ionic species used to draw the diagram are the same as those used to define the protection potential (2.74), the diagram shown allows the ***stability domains*** of the metal, of the oxide and of the Al^{3+} and $Al(OH)_4^-$ ions to be identified.

The line representing reaction (1) of Table 2.17 indicates the protection potential of aluminum. At lower potentials, the rate of corrosion of aluminum metal is negligible. The lines that correspond to reactions (3) and (4) of Table 2.17 define the limits of stability for gibbsite as a function of pH.

Stability domain of water

Figure 2.18 also shows the stability domain for water. The dotted lines show the reversible potential of the hydrogen electrode (line a),

© 2007, first edition, EPFL Press

$$2\,H^+ + 2\,e = H_2 \tag{2.42}$$

$$E_{rev} = -\,0.059\ pH \tag{2.103}$$

and the oxygen electrode (line b),

$$O_2 + 4\,H^+ + 4\,e = 2\,H_2O \tag{2.88}$$

$$E_{rev} = 1.23 - 0.059\ pH \tag{2.104}$$

under the conditions $P_{H2} = 1$ atm and $P_{O2} = 1$ atm. The zone between the lines (a) and (b) corresponds to the domain where, from the thermodynamic perspective, water is stable at ambient pressure. The data of Figure 2.18 show that, for any given pH, the potential of the hydrogen electrode (line a) is higher than the potential of reactions (1) or (5), which describe the equilibrium between aluminum and the dissolved species Al^{3+} and $Al(OH)_4^-$. Outside the stability domain of the oxide $Al(OH)_3$, aluminum can therefore corrode according to one of the following overall reactions.
In acidic environments:

$$2\,Al + 6\,H^+ = 2\,Al^{3+} + 3\,H_2 \tag{2.105}$$

In alkaline environments:

$$2\,Al + 6\,H_2O + 2\,OH^- = 2\,Al(OH)_4^- + 3\,H_2 \tag{2.106}$$

In general, if, for a given pH, the reversible potential of a metal is below the reversible potential of the hydrogen electrode, the metal can corrode by reaction with protons with the production of hydrogen. Keep in mind, however, that thermodynamic considerations are only valid for the prediction of equilibria and offer no information about the rate at which equilibrium is reached. Potential-pH diagrams tell us about the possibility of corrosion of a metal as a function of potential and pH, but not about the rate of corrosion.

Potential–pH diagram of iron

In view of the importance of steel in construction, the corrosion behavior of iron is of particular importance. Figure 2.19 shows the potential-pH diagram for iron, determined as described above for aluminum. Two oxides are taken into account: hydrated Fe_2O_3 and Fe_3O_4 (magnetite). The concentration of the dissolved species is set to 10^{-6} mol l^{-1}. In Figure 2.19, it can be seen that iron can react with protons in acidic and neutral environments, with the giving off of hydrogen. In basic solution, on the other hand, it is resistant to corrosion because iron oxides do not easily dissolve by reaction with hydroxyl ions. This corresponds to what is observed in practice: in acidic and neutral environments, unprotected iron corrodes easily, while in basic environments, such as concrete, is resists well.

© 2007, first edition, EPFL Press

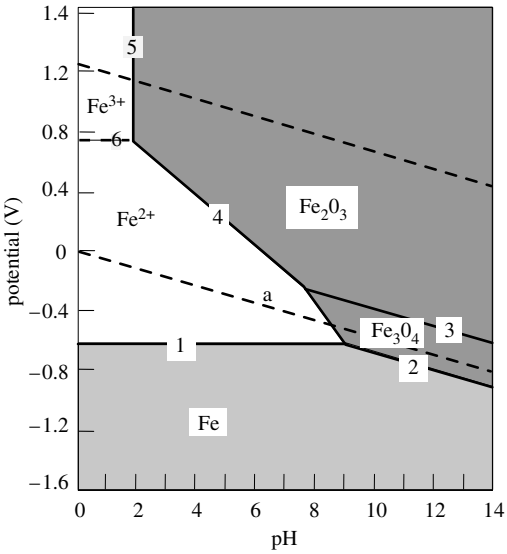


Figure 2.19 Potential-pH diagram of iron (T = 25 °C). The concentrations of the dissolved ionic species, Fe^{2+} and Fe^{3+}, are equal to 10^{-6} mol l^{-1}.

2.3.6 Reference electrodes

To measure the potential of an electrode, a reference electrode is needed. Reference electrodes should satisfy the following set of properties:

- reproducible potential value;
- thermodynamically well–defined reaction;
- non–polarisable (potential not sensitive to current flow);
- ease of use.

In the following we describe briefly the most important reference electrodes used in practice.

Hydrogen electrode

The hydrogen electrode corresponds to the reaction:

$$2\,H^+ + 2\,e = H_2 \tag{2.42}$$

with the reversible potential:

$$E_{rev} = \frac{RT}{2\,F}\ln\frac{a_{H^+}^2}{P_{H_2}} \tag{2.107}$$

At 25 °C and at a partial pressure of hydrogen of 1 bar:

$$E_{rev} = -\,0.059\ \text{pH} \tag{2.103}$$

© 2007, first edition, EPFL Press

Of limited practicality and rarely ever used as a reference electrode, it can be fabricated in the laboratory by bubbling gaseous hydrogen onto an activated platinum electrode (for example, consisting of platinum covered by a electrolytically deposited porous platinum film). Another configuration, more easily built, consists of an electrolysis cell that produces hydrogen on a non–polarizable palladium electrode. When use of the latter is limited to low current density, it establishes a potential that approximately corresponds to the reversible potential of the hydrogen electrode.

Calomel electrode

This electrode is often used in the laboratory as a reference electrode.

$$Hg_2Cl_2 + 2\,e = 2\,Hg + 2\,Cl^- \qquad E^\circ = 0.268\ V \tag{2.108}$$

The Nernst equation for this electrode reaction gives ($T = 25$ °C):

$$E_{rev} = 0.268 - 0.059 \log a_{Cl-} \tag{2.109}$$

According to this equation the reversible potential of the calomel electrode depends only on the activity of the chloride ion in the electrolyte. An electrode whose reversible potential depends uniquely on the activity of anions present in the electrolyte is called ***electrode of the second kind***, or electrode reversible to anions. This behavior implies the presence, at the electrode surface, of a poorly soluble ionic compound. In the case of the calomel electrode, the species in question is the salt Hg_2Cl_2 ("calomel"),

$$Hg_2Cl_2 = Hg_2^{2+} + 2\,Cl^- \tag{2.110}$$

The solubility constant in water has the value:

$$K = a_{Hg_2^{2+}} \cdot a_{Cl^-}^2 = 6.42 \times 10^{-18}\ \text{mol l}^{-1} \tag{2.111}$$

To obtain the relationship between the standard potential of the electrode reaction (2.108) and that of the mercury electrode,

$$Hg_2^{2+} + 2\,e = 2\,Hg \qquad E^\circ = 0.796\ V \tag{2.112}$$

we proceed as for electrode reactions involving the formation of a complex. By setting

$$E_{rev} = 0.268 - 0.059 \log a_{Cl^-} = 0.796 + \frac{0.059}{2} \log a_{Hg_2^{2+}} \tag{2.113}$$

we thus obtain a relationship between the difference of the standard potential and the ***solubility constant***:

$$\Delta E^\circ = 0.796 - 0.268 = \frac{0.059}{2} \log K \tag{2.114}$$

© 2007, first edition, EPFL Press

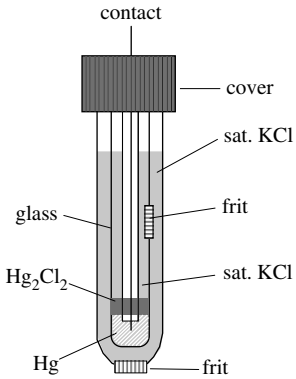


Figure 2.20 The calomel reference electrode.

Figure 2.20 schematically shows one type of calomel electrode that is commonly employed. The electrolyte is a saturated solution of KCl. The potential of this electrode, called the ***saturated calomel electrode,*** abbreviated *sce*, has the value of

$$E_{\text{rev}} = 0.241 \text{ V}$$

relative to the standard hydrogen electrode.

Other reference electrodes

Table 2.21 gives the potentials for different reference electrodes. With the exception of the copper sulfate electrode, all are electrodes of the second kind. Equations (2.105), as well as (2.111) to (2.114), give the corresponding half–cell reaction.

Table 2.21 Potential of typical reference electrodes (25°C)

Electrode	Electrolyte	E / V
Calomel	KCl, saturated	0.241
	KCl, 1 M	0.280
	KCl, 0.1 M	0.333
Mercurous sulfate	K_2SO_4, saturated	0.658
Mercuric oxide	NaOH, 1 M	0.098
Silver chloride	KCl, saturated	0.195
Copper sulfate	$CuSO_4$, saturated	0.316

$$Hg_2SO_4 + 2\,e = 2\,Hg + 2\,SO_4^{2+} \qquad E^\circ = 0.613 \text{ V} \tag{2.115}$$

$$HgO + H_2O + 2\,e = Hg + 2\,OH^- \qquad E^\circ = 0.098 \text{ V} \tag{2.116}$$

© 2007, first edition, EPFL Press

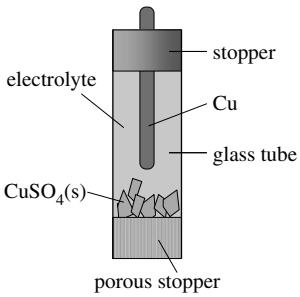


Figure 2.22 The copper sulfate reference electrode.

$$AgCl + e = Ag^+ + Cl^- \qquad E° = 0.222\ V \qquad (2.117)$$

$$Cu^{2+} + 2\,e = Cu \qquad E° = 0.340\ V \qquad (2.118)$$

With an external configuration identical to the calomel electrode, the ***mercurous sulfate*** electrode allows in certain environments to exclude any electrolyte contamination by chloride ions.

The ***mercuric oxide*** electrode is useful for measurements in strongly alkaline environments.

The ***silver chloride*** electrode is easily made by anodic polarization of a silver wire in a concentrated chloride solution, but many different commercial designs are also available. Because there is a risk of mercury pollution during the disposal of used calomel electrodes, these are increasingly replaced by silver chloride electrodes.

The potential of the ***copper sulfate*** electrode (Figure 2.22), particularly inexpensive and easy to build, is less stable than the other reference electrodes mentioned here. It is mostly used to measure the potential of soils.

2.4 A CLOSER LOOK AT ELECTROCHEMICAL CELLS

2.4.1 Electrochemical potential

The chemical potential μ_i describes the energetic state of a *neutral* species. Its value depends on the pressure, the temperature, and the chemical composition of the phase under consideration. The energetic state of a *charged* species (ions, electrons) depends in addition on the electric potential of the phase Φ, called the ***Galvani potential*** (Φ_m refers to the electric potential of a metallic phase, Φ_σ to that of an electrolyte). The ***electrochemical potential*** μ_i^*, defined by (2.119), describes the energetic state of charged species.

$$\mu_i^* = \mu_i + z_i\, F\, \Phi \qquad (2.119)$$

© 2007, first edition, EPFL Press

Here, z_i designates the charge of the species i. For neutral species, $z_i = 0$ and therefore $\mu_i^* = \mu_i$.

More precisely, the electrochemical potential represents the partial free energy of a charged species. The separation into a chemical part (μ_i) and an electrical part ($z_i\ F\Phi$) does not, in general, correspond to any physical reality. Nevertheless, the formalism of (2.119) proves to be quite useful for the analysis of equilibria that involve ions and electrons. The electrochemical potential plays the role here that the chemical potential plays in equilibria between neutral species. The condition for electrochemical equilibrium is this written as:

$$0 = \sum \upsilon_i\ \mu_i^* \tag{2.120}$$

Care must be taken not to confuse the electrochemical potential, which is given in units of J/mole (as for the chemical potential) with the electrode potential, measured in volts.

Electric potential between two phases of the same composition

Between two phases of the same composition, designated by $'$ and $''$, the chemical potential of a species i is identical, in contrast to the electrochemical potential which can vary. According to (2.119), the difference is equal to:

$$\mu_i^{*\prime\prime} - \mu_i^{*\prime} = z_i\ F\,(\Phi'' - \Phi') \tag{2.121}$$

Two copper wires, designated M″ and M′, are connected to a voltmeter. Because the chemical potential of the electrons in the two wires are identical, the measured voltage gives the difference in their electrochemical potentials.

$$U = \Phi_{M''} - \Phi_{M'} = -\frac{\mu^*_{e,M''} - \mu^*_{e,M'}}{F} \tag{2.122}$$

2.4.2 Single–electrolyte electrochemical cells

In electrochemical cells without liquid junctions, the two electrodes are in contact with the same electrolyte of uniform concentration. For example, the cell shown in Figure 2.23 is made of a lead electrode and a lead amalgam electrode (lead dissolved in mercury), in contact with an aqueous solution of $PbCl_2$. This cell corresponds to the schematic representation (2.123), where M′ and M″ refer to the metal of the two conductors attached to the voltmeter.

$$M' \mid Pb_{(s)} \mid PbCl_2, H_2O \mid Pb(Hg) \mid M'' \tag{2.123}$$

The overall reaction is given by the transfer of one mole of lead from the lead electrode to the lead amalgam electrode.

$$Pb_{(s)} \rightarrow Pb_{(Hg)} \tag{2.124}$$

© 2007, first edition, EPFL Press

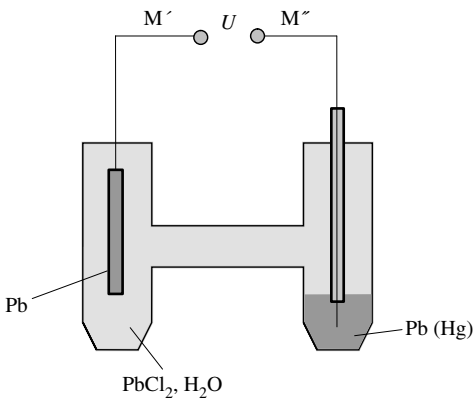


Figure 2.23 The lead/ lead amalgam cell.

The free energy of reaction is written as:

$$\Delta G = \mu_{\mathrm{Pb(Hg)}} - \mu_{\mathrm{Pb(s)}} = RT \ln a_{\mathrm{Pb(Hg)}} \tag{2.125}$$

The reversible potential therefore has the value:

$$E_{\mathrm{rev}} = -\frac{RT}{nF} \ln a_{\mathrm{Pb(Hg)}} \tag{2.126}$$

The two partial reactions, anodic and cathodic, are:

$$\mathrm{Pb}^{2+} + 2e = \mathrm{Pb}_{\mathrm{(s)}} \qquad \text{(anode)} \tag{2.127}$$

$$\mathrm{Pb}^{2+} + 2e = \mathrm{Pb}_{\mathrm{(Hg)}} \qquad \text{(cathode)} \tag{2.128}$$

The reaction (2.124) implies that there are two electrons transferred per lead atom (n=2). According to (2.122), the voltage at the terminals is equal to the difference in the electrochemical potentials of the electrons of the two metal conductors M″ and M′. In order to calculate U, one must write out the different equilibria of the cell:

- equilibria at the electrode–electrolyte interfaces:

$$\text{anode:} \quad 2\mu^*_{\mathrm{e,Pb(s)}} + \mu^*_{\mathrm{Pb^{2+}}} - \mu_{\mathrm{Pb(s)}} = 0 \tag{2.129}$$

$$\text{cathode:} \quad \mu_{\mathrm{Pb(Hg)}} - 2\mu^*_{\mathrm{e,Pb(Hg)}} - \mu^*_{\mathrm{Pb^{2+}}} = 0 \tag{2.130}$$

- equilibria between the electrons at the electrodes and those of the conductors:

$$\text{lead-M}'': \qquad \mu^*_{\mathrm{e,Pb(s)}} = \mu^*_{\mathrm{e,M'}} \tag{2.131}$$

© 2007, first edition, EPFL Press

$$\text{amalgam-M}': \quad \mu^*_{\text{e,Pb(Hg)}} = \mu^*_{\text{e,M}''} \tag{2.132}$$

By combining these equations, the electrochemical potential of the Pb^{2+} ion can be eliminated. After rearranging one gets:

$$\mu^*_{\text{e,M}''} - \mu^*_{\text{e,M}'} = \frac{1}{2}\left(\mu_{\text{Pb(Hg)}} - \mu_{\text{Pb(s)}}\right) = -F\,(\Phi_{\text{M}''} - \Phi_{\text{M}'}) \tag{2.133}$$

$$U = \Phi_{\text{M}''} - \Phi_{\text{M}'} = -\frac{RT}{2F}\ln a_{\text{Pb(Hg)}} \tag{2.134}$$

We deduce from (2.134) and (2.126) that the voltage measured at the terminals is equal to the reversible potential of the overall reaction (2.124). This reasoning is valid for all single electrolyte cells at equilibrium. The cell (2.113) can be used to determine experimentally the activity of lead in mercury.

2.4.3 Electrochemical cells with a liquid junction

Liquid junction electrochemical cells are the most common electrochemical cells. They include, by definition, two electrolytes. The liquid junction refers to the contact region between the two electrolytes.

Example of a liquid junction cell: The Daniell cell

This cell (Figure 2.10) is made of a copper electrode immersed in a solution of copper sulfate, and a zinc electrode immersed in a solution of zinc sulfate. It is schematically represented by (2.135), with a double vertical line to show the liquid junction.

$$\text{Cu}' \mid \text{Zn} \mid \text{ZnSO}_4, \text{H}_2\text{O} \parallel \text{CuSO}_4, \text{H}_2\text{O} \mid \text{Cu}'' \tag{2.135}$$

A porous element, such as a glass frit, separates the two electrolytes. It prevents the two solutions from mixing while allowing the ions to migrate from one solution to the other, thereby providing the means for an ionic current to flow between the two electrodes.

At the interface of the two electrolytes, there exists no real equilibrium situation, because of the presence of strong concentration gradients, which drive ion diffusion across the junction. After an infinitely long period of time, the concentrations would become uniform from one part of the junction to another. However, the presence of a frit or membrane slows down the diffusion phenomena enough to keep the liquid junction in place during a long time. Despite this, the inherent instability of liquid junctions remains the main source of difficulty for those who attempt to rigorously describe the thermodynamics of electrochemical cells.

© 2007, first edition, EPFL Press

Junction potential

As soon as there is a difference of concentration between two electrolytes separated by a frit, the ions will attempt to diffuse through the frit in order to reduce the difference in chemical potential and to allow the system to approach equilibrium. However, certain ions are capable of diffusing more rapidly than others. The result is the formation of a charge imbalance that creates an electric field across the junction and, thus, an electric potential difference between the two electrolytes: the ***junction potential***, $\Delta\Phi_{JC}$. The resulting electric field in the junction slows down the more rapid ions and accelerates the slower ones so that the electrolyte remains electrically neutral everywhere.

The junction potential changes the terminal voltage of the cell:

$$U = E_{rev} + \Delta\Phi_{JC} \tag{2.136}$$

The reversible potential $E_{rev} = -\Delta G/2F$ refers to the overall reaction of the cell, for the Daniell cell:

$$Zn + CuSO_4 = Cu + ZnSO_4 \tag{2.137}$$

The junction potential is equal to

$$\Delta\Phi_{JC} = \Phi_\sigma^{II} - \Phi_\sigma^{I} \tag{2.138}$$

where Φ_σ^{II} and Φ_σ^{I} represent the Galvani potential of the two solutions near the liquid junction (for the Daniell cell, the superscript II represents the $CuSO_4$ solution and I the $ZnSO_4$ solution). For most electrochemical cells, the junction potential is only on the order of a few mV. It increases with the difference in either ion concentration or ion mobility of the two solutions.

It is only rarely possible to calculate $\Delta\Phi_{JC}$ with precision, and in general one must be satisfied with an approximation, such as that given by the ***Henderson equation*** (2.139) [8].

$$\Delta\Phi_{JC} = -\frac{RT}{F}\frac{\sum z_i u_i (c_i^{II} - c_i^{I})}{\sum z_i^2 u_i (c_i^{II} - c_i^{I})} \cdot \ln \frac{\sum z_i^2 u_i c_i^{II}}{\sum z_i^2 u_i c_i^{I}} \tag{2.139}$$

Here, u_i designates the ion mobility of species *i,* and the superscripts ^{II} and ^{I} indicate the electrolyte phases of the two sides of the junction.

Figure 2.24 shows (a) the liquid junction between two solutions of HCl of different concentrations and (b) the liquid junction between a solution of HCl and a solution of NaCl of identical concentration.

HCl diffuses across the junction (a). The mobility of protons is significantly higher than that of Cl^- ions. The concentration of protons therefore tends to build up more quickly than that of Cl^- ions on the 0.1 M side of the junction, resulting in a situation where the potential of the more dilute solution becomes the positive with respect to the concentrated solution. The potential difference slows down the

© 2007, first edition, EPFL Press

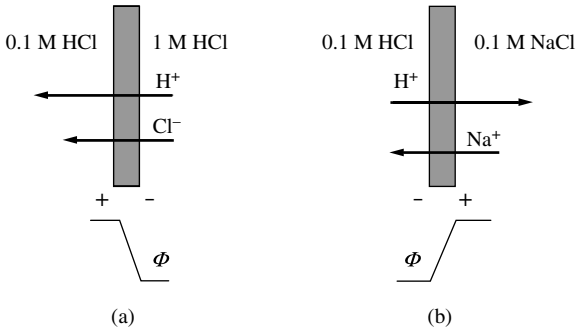


Figure 2.24 The origin of the junction potential: (a) upon contact between two HCl solutions, 0.1M and 1M; and (b) upon contact of a 0.1M HCl solution with a solution of 0.1M NaCl.

movement of the H^+ (repulsion between charges of the same sign) and accelerates that of the Cl^- (attraction between charges of opposite sign).

At junction (b), Na^+ and H^+ ions diffuse in opposite directions. The higher mobility of the H^+ ions increases the potential of the NaCl solution relative to that of the HCl solution. This slows down the movement of the protons towards the NaCl solution and accelerates that of Na^+ towards the solution of HCl.

Concentration cells

These are liquid–junction cells consisting of two identical electrodes, each in contact with a solution of same chemical composition, but of differing concentration. Concentration cells not only play a role in localized corrosion (see Chapter 7), but they represent also one of the rare instances where the junction potential can be calculated with precision. This calculation offers useful insight into the origin of junction potentials.

The electrochemical cell (2.140), shown in Figure 2.25, will provide our example.

$$\mathrm{Cu'} \mid \mathrm{CuSO_4}(c_1), \mathrm{H_2O} \parallel \mathrm{CuSO_4}(c_2), \mathrm{H_2O} \mid \mathrm{Cu''} \tag{2.140}$$

Here c_1 and c_2 refer to the copper sulfate concentrations. We assume that $c_2 > c_1$. If a current flows, we observe a transfer of copper sulfate from the more concentrated to the more dilute solution. The overall reaction of the cell is given by (2.141).

$$\mathrm{Cu'} + (1 - t_+)\,\mathrm{CuSO_4}(c_2) \rightarrow \mathrm{Cu''} + (1 - t_+)\,\mathrm{CuSO_4}(c_1) \tag{2.141}$$

By the time a total charge of $2F$ is transferred from the anode to the cathode, one mole of Cu has dissolved at the anode and an identical quantity of Cu has deposited at the cathode. And yet, only a fraction t_+ of the Cu^{2+} ions liberated at the anode cross the junction because, in the electrolyte, both cations and anions contribute to the ionic current. The Cu^{2+} ions carry a fraction t_+ of the total current, with the SO_4^{2-} ions providing the complementary fraction of the current $t_- = 1 - t_+$ (we assume that there

© 2007, first edition, EPFL Press

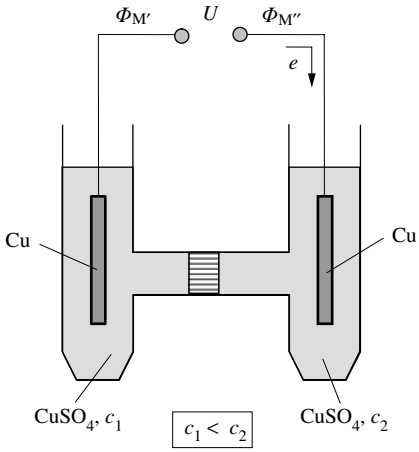


Figure 2.25 A concentration cell.

are no other ions in solution). A charge balance shows that the anode compartment obtains a net gain of $(1-t_+)$ moles of $CuSO_4$.

Table 2.26 Charge balance for the anodic and cathodic compartments of the concentration cell (2.140).

	Anode compartment (concentration c_1)	Cathode compartment (concentration c_2)
By reaction at the anode:	+ 1 mole Cu^{2+}	0
By reaction at the cathode:	0	– 1 mole Cu^{2+}
By migration across the junction:	$-t_+$ mole Cu^{2+}	$+t_+$ mole Cu^{2+}
	$+(1-t_+)$ mole SO_4^{2-}	$-(1-t_+)$ mole SO_4^{2-}
Total:	$+(1-t_+)$ mole $CuSO_4$	$-(1-t_+)$ mole $CuSO_4$

The free energy of reaction corresponding to the transfer of $(1-t_+)$ moles of copper sulfate, from the solution of concentration $c_{(2)}$ to the solution of concentration $c_{(1)}$, is equal to ($f_{(1)}$ and $f_{(2)}$ are activity coefficients):

$$\Delta G = (1-t_+)RT \ln f_{(1)}c_{(1)} - (1-t_+)RT \ln f_{(2)}c_{(2)} \tag{2.142}$$

which gives for the cell potential:

$$E_{\text{rev,cell}} = -(1-t_+)\frac{RT}{2F} \ln \frac{a_{(1)}}{a_{(2)}} \tag{2.143}$$

This value corresponds to the voltage measured at the terminals of the cell, U.

$$E_{\text{rev,cell}} = U \tag{2.144}$$

© 2007, first edition, EPFL Press

The difference of the reversible potentials of the electrode reactions, taken individually

$$Cu^{2+}_{(c_1)} + 2e = Cu \tag{2.145}$$

$$Cu^{2+}_{(c_2)} + 2e = Cu \tag{2.146}$$

in turn are equal to:

$$\Delta E_{\text{rev}(2,1)} = E_{\text{rev}(c_2)} - E_{\text{rev}(c_1)} = \frac{RT}{2F} \ln \frac{a_{(2)}}{a_{(1)}} \tag{2.147}$$

where $a_{(1)}$ and $a_{(2)}$ represent the $CuSO_4$ activity in the two half cells.

The difference between $U = E_{\text{rev,cell}}$ and $\Delta E_{\text{rev}(2,1)}$ corresponds to the value of the junction potential:

$$\Delta\Phi_{\text{JC}} = U - \Delta E_{\text{rev}(2,1)} \tag{2.148}$$

With (2.143) and (2.147) we obtain:

$$\Delta\Phi_{\text{JC}} = -t_+ \frac{RT}{2F} \ln \frac{a_{(2)}}{a_{(1)}} \tag{2.149}$$

The study of concentration cells allows us to determine, among other parameters, the values for t_+ and t_-, the ***transport numbers*** for the cation and the anion, respectively. These parameters will be discussed in more detail in Section 4.3.

2.5 PHYSICAL POTENTIAL SCALE

2.5.1 Reversible potential and the Fermi level

In metals, the conduction band overlaps with the valence band. It is only partially filled with electrons. The Fermi distribution function describes the distribution of electrons over the number of energetically accessible states. The probability of occupation, f_x, of a given energy level E_x is given by

$$f_x = \frac{1}{1 + \exp\left[(E_x - E_F)/kT\right]} \tag{2.150}$$

where E_F is called the Fermi energy or the ***Fermi level***. The Fermi energy is measured relative to the energy of a free electron at rest in a vacuum, whose value is assumed to be exactly zero. At absolute zero, all the energy levels below E_F are occupied. At higher temperatures, E_F corresponds to an occupation probability of 1/2.

Quite generally, the Fermi level characterizes the energy of electrons in a solid. For metals, the Fermi energy of electrons is equal to their electrochemical potential.

© 2007, first edition, EPFL Press

$$E_F = \mu_e^* \tag{2.151}$$

The voltage U measured between two conductors M′ and M″ of the same chemical composition corresponds to the difference in Fermi energies of the electrons.

$$U = \Phi_{M''} - \Phi_{M'} = -\frac{1}{F}(\mu_e^{*''} - \mu_e^{*'}) = -(E_F'' - E_F')\frac{1}{F} \tag{2.152}$$

In order to relate electrochemical behavior to the electronic properties of a metal or of a semiconductor, one must find a relationship between the Fermi level and the reversible potential. In the next paragraph, we develop such a relationship for metals, starting with the half–cell reaction (2.153).

$$\nu_{ox} B_{ox} + n\, e = \nu_{red} B_{red} \tag{2.153}$$

The subscripts *ox* and *red* here refer to oxidized and reduced species in solution, in equilibrium with an inert electrode, for example Fe^{3+} and Fe^{2+} ions in contact with a platinum electrode. At equilibrium, with (2.120):

$$\nu_{red}\, \mu_{red}^* - n\, \mu_e^* - \nu_{ox}\, \mu_{ox}^* = 0 \tag{2.154}$$

The electrochemical potential of the electrons in the metal is therefore equal to

$$\mu_e^* = \frac{1}{n}(\nu_{red}\, \mu_{red}^* - \nu_{ox}\, \mu_{ox}^*)$$

$$= \frac{1}{n}[(\nu_{red}\, \mu_{red} - \nu_{ox}\, \mu_{ox}) + F\, \Phi_\sigma\, (z_{red}\, \nu_{red} - z_{ox}\, \nu_{ox})] \tag{2.155}$$

Here, Φ_σ represents the electric potential (Galvani potential) in the solution. The law of conservation of charge implies that:

$$z_{red}\, \nu_{red} - z_{ox}\, \nu_{ox} = -n \tag{2.156}$$

By the substitution $\mu_i = \mu_i^o + RT\, ln\, a_i$, followed by a rearrangement:

$$\mu_e^* = \frac{1}{n}\left[\Delta G^\circ + RT\, \ln \frac{a_{red}^{\nu_{red}}}{a_{ox}^{\nu_{ox}}}\right] - F\, \Phi_\sigma \tag{2.157}$$

The term in parentheses corresponds to the free energy of reaction. With (2.36), we get (2.158).

$$\mu_e^* = -F\, E'_{rev} - F\, \Phi_\sigma \tag{2.158}$$

Here E'_{rev} is the reversible potential of the half cell measured with respect to an arbitrary reference. Normally, we measure the reversible potential of half cells by using the standard hydrogen electrode as the point of reference. Equation (2.158) then becomes:

© 2007, first edition, EPFL Press

$$\mu_e^* = -F\,E_{rev} - F\,\Phi_\sigma^\circ \tag{2.159}$$

This equation shows us that the reversible potential of an electrode reaction corresponds to the electrochemical potential (Fermi energy) of the electrons in the metal, with the presence of an additional constant.

$$E_{rev} = -\frac{\mu_e^*}{F} - \Phi_\sigma^\circ = -\frac{E_F}{F} - \Phi_\sigma^\circ \tag{2.160}$$

Unfortunately, the value of Φ_σ° is not directly obtainable because, in order to measure the potential of an electrode, one must use an electrochemical cell of two electrodes, and the terms $F\Phi_\sigma^\circ$then cancel each other out (subsection 2.2.2). For example, the terminal voltage for the single–electrolyte cell (2.161),

$$M' \mid B_{ox}, B_{red}, H^+, H_2 \mid M'' \tag{2.161}$$

consisting of the half cell (2.153) and the standard hydrogen electrode, corresponds to the difference of the Fermi levels in the metals M″ and M′. It is therefore independent of Φ_σ°.

$$U = -\frac{\mu_e^{*\prime\prime} - \mu_e^{*\prime}}{F} = -\frac{E_F'' - E_F'}{F} \tag{2.162}$$

In order to obtain the value of Φ_σ°, another method must be employed.

2.5.2 Relation between electrochemical and physical potential scales

In electrochemistry, we indicate the half–cell potentials relative to that of the standard hydrogen electrode.

$$H^+_{(aq)} + e_{(m)} = \frac{1}{2}\,H_{2(g)} \tag{2.163}$$

In this equation $H^+_{(aq)}$ designates the solvated proton, $e_{(m)}$ represents an electron in the metal, and $H_{2(g)}$ is gaseous hydrogen. According to (2.120), the equilibrium condition at the metal–electrolyte–gas interface for the reaction (2.163) is given by:

$$\frac{1}{2}\mu^\circ_{H2(g)} - \mu^*_{H+(aq)} - \mu^*_{e(m)} = 0 \tag{2.164}$$

At standard state, according to (2.45), $\mu_{H+(aq)} = \mu^\circ_{H+(aq)} = 0$ and $\mu^\circ_{H2(g)} = 0$. By substituting $\mu^*_{H+(aq)} = \mu^\circ_{H+(aq)} + F\,\Phi_\sigma^\circ = F\,\Phi_\sigma^\circ$ we then obtain:

$$\mu^*_{e(m)} = -F\,\Phi_\sigma^\circ \tag{2.165}$$

Comparison with (2.164) shows that if we manage to determine $\mu^*_{H+(aq)}$ $-1/2\,\mu^\circ_{H2(g)}$, we can calculate $\mu^*_{e(m)}$ and, by extension, $F\Phi_\sigma^\circ$.

To these ends, we break down the half–cell reaction into several elementary steps, according to the following reaction scheme [9]:

© 2007, first edition, EPFL Press

$$H^+_{(aq)} \rightarrow H^+_{(g)} \qquad \Delta G^\circ = 1088 \text{ kJ mole}^{-1}$$

$$e_{(m)} \rightarrow e_{(g)} \qquad \Delta G^\circ = -\mu^*_{e(m)}$$

$$H^+_{(g)} + e_{(g)} \rightarrow H_{(g)} \qquad \Delta G^\circ = -1313 \text{ kJ mole}^{-1}$$

$$H_{(g)} \rightarrow \frac{1}{2} H_{2(g)} \qquad \Delta G^\circ = -203 \text{ kJ mole}^{-1}$$

$$H^+_{(aq)} + e_{(m)} \rightarrow \frac{1}{2} H_{2(g)} \qquad \Delta G^\circ = -428 \text{ kJ mole}^{-1} - \mu^*_{e(m)} = 0 \tag{2.166}$$

The solvated proton is transformed into an unsolvated gaseous proton. 1088 kJ mol^{-1} is required for the free energy of solvation. The electrochemical potential of the electron $\mu^*_{e(m)}$ is equivalent to the work function, in other words to the work to be supplied to transfer an electron from the Fermi level within the metal to the vacuum level. The reaction between the proton and the electron yields atomic hydrogen, generating 1313 kJ mol^{-1}. This same value, but of opposite sign, corresponds to the ionization energy of atomic hydrogen. Finally, two hydrogen atoms react together to form a molecule. The free energy of reaction per mole of atomic hydrogen releases 203 kJ mol^{-1}.

The sum of these values yields the free energy of reaction for the standard hydrogen electrode. By definition, E° and therefore ΔG° are zero; thus the electrochemical potential of the electrons (equal to the Fermi energy) is equal to:

$$\mu^*_{e(m)} = -F \cdot \Phi^\circ_\sigma = -428 \text{ kJ mole}^{-1} \tag{2.167}$$

The Fermi energy is generally expressed in eV: 1 eV = 96.485 kJ mole^{-1}.

$$E_F = -F \cdot \Phi^\circ_\sigma = -4.44 \text{ eV} \tag{2.168}$$

With (2.159), it is possible to define an energy scale of the electrons that corresponds to the scale of the standard potentials of electrode reactions.

$$E_F = -F \cdot E^\circ - 4.44 \text{ eV} \tag{2.169}$$

Figure 2.27 compares the standard potentials measured in V relative to the standard hydrogen electrode, ***electrochemical scale***, with the corresponding Fermi energies of the electrons in the metal measured in eV relative to the vacuum level, the ***physical scale***. Dividing (2.167) by $F = 96485$ C mol^{-1} we obtain $\Phi^\circ_\sigma = 4.44$ V. The potential of the standard hydrogen electrode therefore has a potential of 4.44 V on the physical scale. The physical scale of potentials is frequently used in the study of the electrochemistry of semiconductors.

© 2007, first edition, EPFL Press

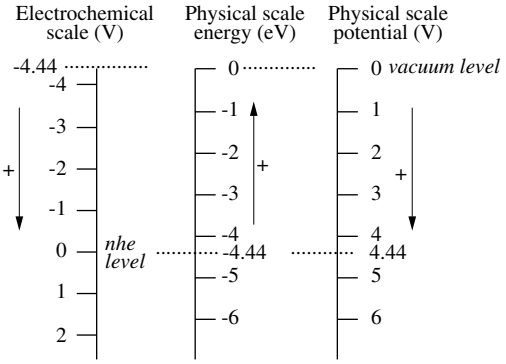


Figure 2.27 Comparison of potential scales based on the standard hydrogen electrode and on the vacuum-level electron energy, respectively.

General bibliography

D. R. Gaskell, *Introduction to Metallurgical Thermodynamics*, 2nd edition, MacGraw Hill International Edition, New York (1981) 611 pp.

R. A. Robinson, R.H. Stokes, *Electrolyte Solutions*, 2nd edition, Butterworths, London (1970) 571 pp.

J. S. Newman, *Electrochemical Systems*, 2nd edition, Prentice Hall, Englewood Cliffs (1991) 560 pp.

References

[1] I. Barin, O. Knacke, *Thermochemical Properties of Inorganic Substances, Volume I,* Supplement, Springer Verlag, Heidelberg (1973, 1977), pp. 1–861.

[2] H. E. Mc Gannon, *The Making, Shaping and Treating of Steel*, 9th edition, US Steel Company, Pittsburgh Pa. (1971) pp. 297–98.

[3] G. Kortüm, *Lehrbuch der Elektrochemie*, 5th edition, Verlag Chemie, Weinheim (1972) p. 617.

[4] R. A. Robinson, R. H. Stokes, *Electrolyte Solutions*, 2nd edition, Butterworths, London (1970) pp. 74 and 230.

[5] R. A. Robinson, R. H. Stokes, *Electrolyte Solutions*, 2nd edition, Butterworths, London (1970) pp. 491–503.

[6] A. J. Bard, R. Parsons, J. Jordan, *Standard Potentials in Aqueous Solution*, Marcel Dekker Inc, New York (1985) pp. 1–834.

[7] M. Pourbaix, *Atlas of Electrochemical Equilibria in Aqueous Solutions*, 2nd edition, NACE, Houston Texas (1974) pp. 1–644.

[8] J. S. Newman, *Electrochemical Systems*, 2nd edition, Prentice Hall, Englewood Cliffs (1991) p. 137.

[9] A. J. Bard, R. Parsons, J. Jordan, *Standard Potentials in Aqueous Solution*, Marcel Dekker Inc, New York (1985) p.118.

© 2007, first edition, EPFL Press

CHAPTER 3

SURFACES AND INTERFACES

3.1 SURFACE TENSION EFFECTS

3.1.1 Surface energy

In order to enlarge the surface of a solid or a liquid by the increment dA, a certain number of bulk atoms must be moved towards the surface. Because the surface atoms have fewer bonds than those situated in the bulk, the Gibbs free energy of the solid or liquid increases by an amount dG.

$$\mathrm{d}G = \gamma \cdot \mathrm{d}A \tag{3.1}$$

The definition of the ***surface energy*** γ emerges from a more general formulation of the Helmholtz or Gibbs free energy change of a thermodynamic system in the presence of surface effects:

$$\mathrm{d}F = -P\,\mathrm{dV} - S\,\mathrm{d}T + \Sigma\,\mu_\mathrm{i}\,\mathrm{d}n_\mathrm{i} + \gamma\mathrm{d}A \tag{3.2}$$

$$\mathrm{d}G = V\,\mathrm{d}P - S\,\mathrm{d}T + \Sigma\,\mu_\mathrm{i}\,\mathrm{d}n_\mathrm{i} + \gamma\mathrm{d}A \tag{3.3}$$

$$\gamma = \left(\frac{\partial F}{\partial A}\right)_{\mathrm{V,T,n_i}} = \left(\frac{\partial G}{\partial A}\right)_{\mathrm{P,T,n_i}} \tag{3.4}$$

To get a rough estimate of the surface energy of metals, we calculate the overall change in bond energy as an atom moves from the bulk to the surface is calculated.

In a metal with face-centered cubic (fcc) arrangement, each atom has 12 near neighbors and therefore the same number of bonds. The energy required to break these bonds is equal to the energy of sublimation (on the order of 500 kJ/mole for many metals). Atoms will have 9 near neighbors if they form part of a surface with the orientation corresponding to the Miller indices (111). With an average value of 10^{19} atoms per square meter (0.17×10^{-4} moles), we can make the following estimate for the surface energy:

$$(0.17\times10^{-4})\cdot((12-9)/12)\cdot(500\times10^{3}) \approx 2\ \mathrm{J/m^2} \tag{3.5}$$

This order of magnitude estimate agrees with measured values.

© 2007, first edition, EPFL Press

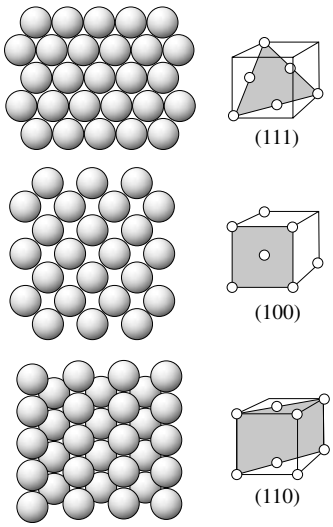


Figure 3.1 Surface atoms of a fcc crystal, with orientations (111), (100), and (110).

Thus, according to the simple model above, the surface energy of metals depends on the crystal orientation. For a metal of structure fcc, atoms belonging to surfaces with the orientations (111), (100), and (110), have, respectively, 9, 8, and 6 near neighbors (Figure 3.1). The surface energy thus increases in the order $\gamma(111) < \gamma(100) < \gamma(110)$. Because surfaces of orientation (111) are most compact, they are also the most stable from a thermodynamic point of view.

Energy of grain boundaries

Grain boundaries are discontinuities in the regular atomic structure of a crystal (Figure 3.2). In order to create a grain boundary, atoms from the bulk have to be displaced towards the boundary, which requires energy. Equation (3.1) can still be used to calculate the change in free energy, but the surface energy γ must be replaced

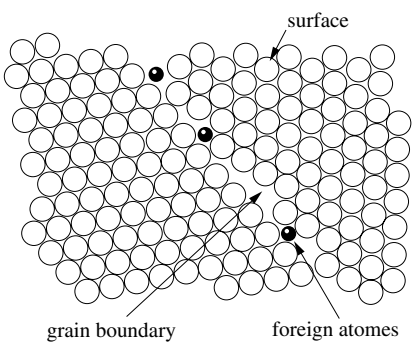


Figure 3.2 Segregation of foreign atoms at a grain boundary.

© 2007, first edition, EPFL Press

by the grain boundary energy γ_{gb}. The latter increases with the misorientation between the adjacent grains, but it always remains less than the surface energy γ.

Because of their higher energy, the atoms located at grain boundaries are often favored as reaction sites during chemical or electrochemical attack of a metal. In metallography, this behavior is used to bring out the grain structure of polycrystalline metals and alloys. From a metallurgical view, the grain boundaries are favorite sites for the segregation of impurities present in the metal and for the nucleation of precipitates (cf. Chapter 7).

Surface tension

The force per unit length corresponding to the surface energy of a solid or liquid body (J m^{-2} = N m^{-1}) is referred to as the ***surface tension***. The concept of surface tension is particularly useful for describing surface effects in liquids. Figure 3.3 shows a soap film of length L that is stretched between two ends of an U-shaped wire. In order to prevent its surface from retracting, a pulling force of $F_s = 2\gamma L$ is required. A factor of 2 appears because the film has two sides exposed to the air. In order to increase the surface of the film by the increment $dA = L\,dz$, the following amount of work has to be provided:

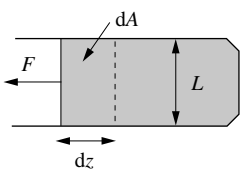


Figure 3.3 Soap film held by U-shaped metal wire.

Table 3.4 Surface tension of different solids and liquids (phase 1) in contact with a vacuum, a gas, or a liquid (phase 2).

Phase 1	Phase 2	T (°C)	γ(Nm^{-1})
Au(s)	vacuum	1000	1.4
Au(l)	He	1200	1.07
Ni(s)	vacuum	1060	2.3
MgO(s)	vacuum	1300	1.2
Hg(l)	vacuum	20	0.472
Hg(l)	water	20	0.375
H_2O	air	25	0.0720
H_2O	air	80	0.0626
Ethanol	air	20	0.0226
Acetone	air	20	0.0237
Glycerol	air	20	0.0634

© 2007, first edition, EPFL Press

$$dW = -F_s \cdot dz = -2\gamma L\, dz \tag{3.6}$$

In reality, the surface tension of a solid or a liquid depends on the phase with which it is in contact as illustrated by the numerical values presented in Table 3.4. More precisely, one should therefore speak of *interfacial tension* rather than surface tension.

3.1.2 Wettability and capillary action

The wettability of a surface indicates the ease with which a liquid drop spreads on a solid surface. It plays an important role in atmospheric corrosion, and in the application of organic coatings such as paints. The wettability is characterized by the contact angle Θ (Figure 3.5), which in turn depends on the three surface tensions γ_{sg}, γ_{sl} and γ_{ls} that act on the line of contact between the solid (s), liquid (l) et gas (g) surfaces. At equilibrium, the sum of the forces parallel to the surface of the solid is equal to zero. This gives the following relationship:

$$\cos\Theta = \frac{\gamma_{gs} - \gamma_{ls}}{\gamma_{gl}} \tag{3.7}$$

The quantity $K_W = \cos\Theta$ is called the ***wetting coefficient***. Perfect wetting corresponds to $K_W = 1$, the absence of wetting results in $K_W \approx -1$.

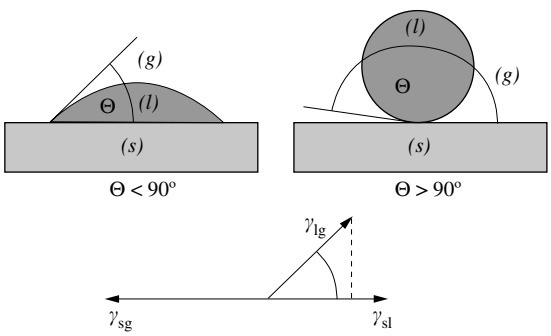


Figure 3.5 Contact angle of a liquid on wetting and non-wetting surfaces.

Young-Laplace Equation

The surface tension induces a pressure difference between the inside and outside of a curved surface. One consequence of this phenomenon is the condition of supersaturation required in order for gas bubbles to form in liquids, as, for example, when hydrogen is released from an electrode. In order to calculate the pressure difference, we determine the variation in Helmholtz free energy when the radius r increases by the increment dr. In the absence of chemical reactions and at a constant temperature, equation (3.2) gives

© 2007, first edition, EPFL Press

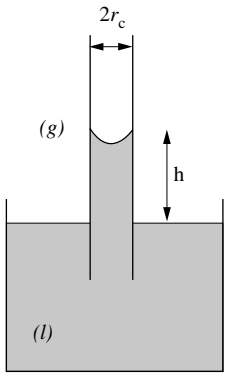


Figure 3.6 Measurement of surface tension by the capillary method

$$dF = -\Delta P\, dV + \gamma dA \tag{3.8}$$

The first term represents the work due to the volume expansion of the gas bubble, for which $\Delta P = P_{int} - P_{ext}$ refers to the pressure difference between inside and outside of the bubble.

For a spherical shape: $dV = 4\pi r^2 dr$, and $dA = 8\pi dr$. With the equilibrium condition $dF = 0$, we obtain the ***Young-Laplace equation***:

$$\Delta P = \frac{2\gamma}{r} \tag{3.9}$$

The surface tension is also responsible for the ***capillary effect***, which causes a liquid to penetrate into a narrow channel against the force of gravity (Figure 3.6). The difference in level between the liquid inside and outside the capillary h corresponds to a pressure difference equal to

$$\Delta P = \frac{2\gamma}{r} = \rho g h \tag{3.10}$$

where ρ is the liquid density and g is the gravitational constant. For a liquid that perfectly wets the wall ($\cos\Theta = 1$), the radius of the capillary, r_c, is equal to the radius of the meniscus, r. As a result:

$$h = \frac{2\gamma}{\rho g r_c} \tag{3.11}$$

With this relationship, the surface tension of a liquid can be evaluated by measuring the height of the liquid level inside the capillary. The book of Adamson (see the bibliography at the end of this chapter) gives an overview of the experimental methods used to measure surface tension for liquids and solids.

© 2007, first edition, EPFL Press

Kelvin equation

The vapor pressure of a small droplet is higher than that of a liquid with a flat surface. For a liquid located in a pore it is even lower. As a consequence water vapor condenses more easily on a porous surface, such as a metal covered with rust, than on a flat surface. This is of importance in atmospheric corrosion as will be discussed in Chapter 8.

Figure 3.7 shows two liquids in equilibrium with a gas phase, namely, a plane surface and a droplet of radius r. The partial vapor pressures are given by P_0 and P', respectively. The transfer of $\mathrm{d}n_\mathrm{i}$ moles of liquid from the plane surface to the drop results in a change of free energy $\mathrm{d}G$.

$$\mathrm{d}G = \gamma \cdot \mathrm{d}A = \gamma \cdot 8\pi r\,\mathrm{d}r \tag{3.12}$$

Now, in each case, the equilibrium implies that the chemical potentials between the gas and the liquid are equal. The transfer of dn_i moles of gas at pressure P_0 to pressure P' therefore results in an free energy change dG identical to the previous change:

$$\mathrm{d}G = \int_{P_0}^{P'} V\,\mathrm{d}P = \mathrm{d}n_i \int_{P_0}^{P'} \frac{RT}{P}\,\mathrm{d}P = \mathrm{d}n_i RT \ln\frac{P'}{P_0} \tag{3.13}$$

With $\mathrm{d}n_\mathrm{i} = (\rho_\mathrm{l} M^{-1})4\pi r^2 \mathrm{d}r$ where ρ_l is the liquid density, we obtain the ***Kelvin equation***:

$$RT \ln\frac{P'}{P_0} = \frac{2\gamma M}{r\rho_1} \tag{3.14}$$

This equation shows that the vapor pressure P' of a droplet depends on its radius and on its surface tension γ. With analogous reasoning, we can obtain an equation for the vapor pressure of a liquid contained in a pore of radius r:

$$RT \ln\frac{P_0}{P'} = \frac{2\gamma M}{r\rho_1} \tag{3.15}$$

In this case P' is, however, smaller than P_0.

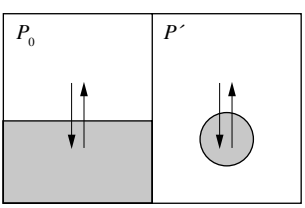


Figure 3.7 Vapor pressure of liquid in equilibrium with a plane surface (P_0) and a droplet (P').

© 2007, first edition, EPFL Press

3.1.3 Gibbs equation

The surface tension of water diminishes when a surfactant, even in small quantities, is added. Similarly, an adsorbed gas reduces the surface energy of a solid. The ***Gibbs equation*** provides a relation between surface tension and surface concentration in multicomponent systems.

Let us suppose that i represents the number of components forming a phase. As before, increasing the surface of the phase by the increment dA amounts to the displacement of a certain number of atoms from the bulk to the surface, resulting in a change dG of the free energy of the system:

$$\mathrm{d}G = \gamma \mathrm{d}A + \sum \mu_\mathrm{i}\, dn_\mathrm{i} \tag{3.16}$$

In this equation, dn_i is negative if atoms are removed from the bulk and placed at the surface. For a given species, the mass balance gives

$$\mathrm{d}n_\mathrm{i} = \Gamma_\mathrm{i} \cdot \mathrm{d}A \tag{3.17}$$

where $\Gamma_\mathrm{i} = \mathrm{d}n_\mathrm{i}/\mathrm{d}A$ represents the ***surface excess*** for the concentration of species i, a concept which is illustrated by Figure 3.8. The surface excess Γ_i corresponds to the difference between the number of moles of i within a volume close to the surface and within an identical volume in the bulk. Because Γ_i is measured in (mol m^{-2}) it is not necessary to know the volume considered.

We define the ***surface free energy*** G_s (J m^{-2}) as the difference between the real free energy of the system, G, and the free energy that the system would have without surface effects, G':

$$G^\mathrm{s} = \frac{G - G'}{A} \tag{3.18}$$

The growth dA of the surface is accompanied by an increase dG of the real free energy of the system:

$$\mathrm{d}G = G^\mathrm{s} \cdot \mathrm{d}A \tag{3.19}$$

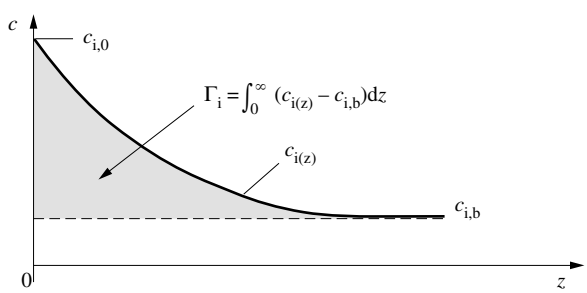


Figure 3.8 Definition of surface excess.

© 2007, first edition, EPFL Press

By comparison with (3.16) and subsequent division by dA, we obtain an equation that relates the surface free energy to the surface tension for a multicomponent system:

$$G^{s} = \gamma + \sum \mu_i \Gamma_i \tag{3.20}$$

On the other hand, if there is only a single component, the second term is set to zero and $G^s = \gamma$. Differentiation of (3.20) gives:

$$dG^{s} = d\gamma + \sum \mu_i \, d\Gamma_i + \sum \Gamma_i \, d\mu_i \tag{3.21}$$

By analogy with bulk phases, the surface free energy depends on T and Γ_i (the influence of pressure being negligible for liquids and solids).

$$dG^{s} = -S^{s} dT + \sum \mu_i \, d\Gamma_i \tag{3.22}$$

For a constant temperature ($dT = 0$), we obtain the ***Gibbs equation*** by comparison with (3.21):

$$d\gamma = -\sum \Gamma_i \, d\mu_i \tag{3.23}$$

For binary systems, to which this equation is usually applied, we get:

$$d\gamma = -\Gamma_1 \, d\mu_1 - \Gamma_2 \, d\mu_2 \tag{3.24}$$

For dilute solutions, we assume that for the solvent $d\mu_1 = 0$.

$$d\gamma = -\Gamma_2 \cdot d\mu_2 \tag{3.25}$$

Since $d\mu_2 = RT \, d\ln a_2$, where a_2 is the activity of the solute, we obtain a relationship that can be used to determine Γ_2:

$$\Gamma_2 = -\frac{1}{RT} \frac{d\gamma}{d\ln a_2} \tag{3.26}$$

According to this last equation, the presence of an excess of solute at the surface results in a reduction of the surface tension. In the same fashion, the adsorption of a gas on a solid causes the surface energy to drop. This is shown in the following equation, valid for an ideal gas of partial pressure (equal to its activity) of P_2.

$$d\gamma = -\Gamma_2 RT \cdot d\ln P_2 \tag{3.27}$$

3.2 ADSORPTION

3.2.1 Heat of adsorption

The adsorption of a gas or a liquid lowers the surface energy. The overall drop of free energy that results is equivalent to the ***free energy of adsorption*** ΔG_{ads}. The Gibbs-Helmholtz equation provides a relationship between the free energy of

© 2007, first edition, EPFL Press

adsorption ΔG_{ads}, the enthalpy of adsorption ΔH_{ads} and the entropy of adsorption ΔS_{ads}:

$$\Delta G_{ads} = \Delta H_{ads} - T\Delta S_{ads} \tag{3.28}$$

For adsorption to take place, ΔG_{ads} must be negative. Now, ΔS_{ads} is always negative, because the adsorbed atoms or molecules lose degrees of freedom in the process. As a result, ΔH_{ads} must be negative, showing that adsorption is an exothermic process. The heat of adsorption Q_{ads}, defined by (3.29), depends on the energies of the bonds formed between the adsorbed atoms and the surface.

$$Q_{ads} = -\Delta H_{ads} \tag{3.29}$$

The value of Q_{ads} allows a distinction to be made between physical and chemical adsorption.

Physical adsorption or physisorption

This phenomenon is due to weak forces, of the Van der Waals type, between the adsorbed species and the surface. It is a reversible phenomenon that does not involve an activation energy. The heat of adsorption does not exceed 50 kJ mol^{-1}, which is comparable to that for the condensation of a gas. Figure 3.9 schematically shows the potential energy of a physisorbed atom or molecule as a function of its distance from the surface. The equilibrium position corresponds to the energy minimum. The adsorbed species must receive energy at least equal to the value of Q_{ads} in order to escape from this position.

During physisorption, several layers of adsorbed atoms may form on top of each other.

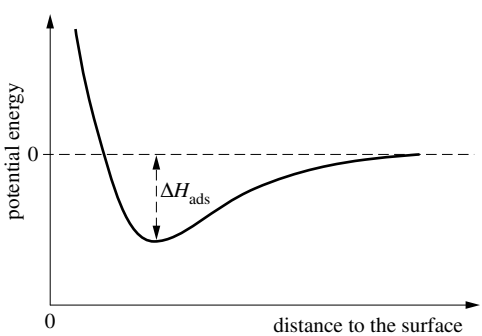


Figure 3.9 Variation of the potential energy with distance in physisorption.

Chemical adsorption or chemisorption

A higher value for the heat of adsorption (from 100 to 500 kJ mol^{-1}) is characteristic of chemical adsorption and implies the presence of strong chemical

© 2007, first edition, EPFL Press

bonds. This process sometimes requires overcoming an activation energy. In general, a single layer of atoms or molecules is adsorbed.

Chemisorption dissociates the molecules of diatomic gases, such as hydrogen or oxygen, transforming them into adsorbed atoms.

$$O_2 \rightleftharpoons 2\,O_{ads} \tag{3.30}$$

$$H_2 \rightleftharpoons 2\,H_{ads} \tag{3.31}$$

This type of reaction is called ***dissociative chemical adsorption***. Figure 3.10 schematically shows the variation in potential energy as a function of the distance from the surface. The first minimum represents the precursor state, corresponding to the physical adsorption of the molecule X_2. The second minimum, clearly much lower, indicates the stable state of chemical adsorption. Even if it can be positive (curve b), the energy of the transition state between the two minima is most often negative (curve a). The activation energy for desorption is then equal to Q_{ads}.

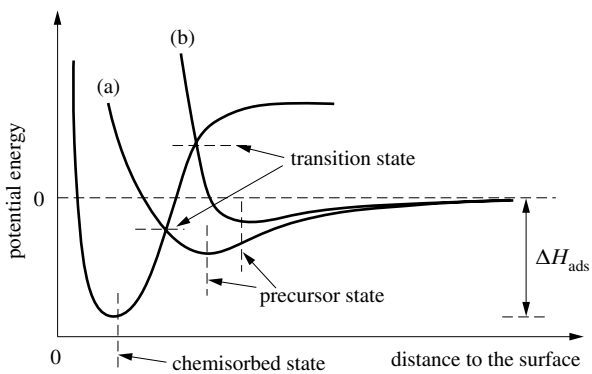


Figure 3.10 Variation of the potential energy with distance during chemisorption. The energy of the transistion state can be either negative (a) or positive (b).

3.2.2 Adsorption isotherms

For a given temperature, the relationship between the amount of adsorbed species and its activity in the gas or liquid phase in contact with the surface is called ***adsorption isotherm***.

The amount of adsorbed species *i* is generally expressed by the ***surface coverage*** θ_i, which describes the fraction of the surface sites occupied by the species:

$$\theta_i = N_i / N_o \tag{3.32}$$

N_i corresponds to the number of sites occupied by the adsorbed species and N_o to the total number of sites per unit of surface.

As we normally associate the activity of a gas with its partial pressure P_i, the adsorption isotherm of a gas is a relationship of the type

© 2007, first edition, EPFL Press

$$\theta_i = f(P_i) \tag{3.33}$$

For a species in solution, such an a corrosion inhibitor, the concentration c_i generally replaces the activity, which gives:

$$\theta_i = f(c_i) \tag{3.34}$$

Langmuir isotherm

Experimental observations have led scientists to develop several theoretical models to describe adsorption isotherms. The best known, due to Langmuir, assumes that there are a fixed number of sites on the surface. Each of these sites is capable of adsorbing a single atom or molecule. In addition, as it is assumed that no interactions between adsorbed species take place, the adsorption energy is constant.

The following derivation is for the Langmuir isotherm of a gas. This is a simple case because the rate of adsorption is proportional to the partial pressure and to the number of non-occupied sites $(1-\theta_i)$.

$$v_{ads} = k_{ads}\ (1 - \theta_i)\ P_i \tag{3.35}$$

Inversely, the desorption rate is proportional to the fraction of sites occupied by the adsorbed gas:

$$v_{des} = k_{des} \cdot \theta_i \tag{3.36}$$

At equilibrium, the two rates are equal:

$$k_{ads}\ (1-\theta_i)\ P_i = k_{des} \cdot \theta_i \tag{3.37}$$

A rearrangement yields the ***surface coverage***:

$$\theta_i = \frac{b_L P_i}{1 + b_L P_i} \tag{3.38}$$

$$b_L = k_{ads}/k_{des} \qquad (m^2\ N^{-1})$$

We call b_L the ***adsorption coefficient***. Figure 3.11 schematically shows the variation in coverage as a function of the product $b_L P_i$. Asymptotically, the curve tends towards $\theta_i = 1$ when $b_L\ P_i >> 1$.

Dissociative chemical adsorption of the molecule X_2 requires two sites:

$$X_2 + 2\ \text{sites} \rightleftharpoons 2\ X_{ads} \tag{3.39}$$

As a result, the equilibrium is written as:

$$k_{ads}\ P_i\ (1-\theta)^2 = k_{des} \cdot \theta^2 \tag{3.40}$$

© 2007, first edition, EPFL Press

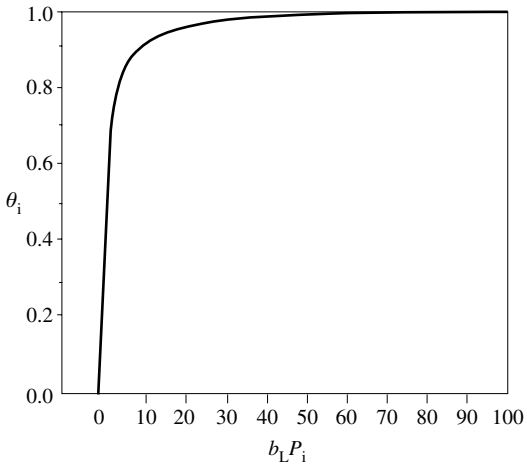


Figure 3.11 Langmuir adsorption isotherm (Equ. (3.38)).

Which gives

$$\theta_i = \frac{(b_L P_i)^{1/2}}{1 + (b_L P_i)^{1/2}} \tag{3.41}$$

Other adsorption isotherms

Not all systems obey the Langmuir isotherm, because certain basic hypotheses of the model may not be valid. It is sometimes observed that the enthalpy of adsorption varies with θ_i. To take this into account, the Temkin isotherm assumes that the enthalpy of adsorption decreases as a function of the coverage. For a gas this results in a relationship of the type:

$$\theta_i = C_{T1} \ln (C_{T2} \cdot P_i) \tag{3.42}$$

where C_{T1} and C_{T2} are constants. The Temkin isotherm is applicable at intermediate levels of coverage: $0.2 < \theta < 0.8$.

*The **Freundlich isotherm*** is an empirical function of the form

$$\theta_i = C_{F1} \cdot P_i^{1/C_{F2}} \tag{3.43}$$

where C_{F1} and C_{F2} are constants.

The ***BET isotherm*** describes the multilayer physical adsorption of a gas [1]. The model is sometimes applied for the measurement of the surface area of porous bodies by determining the amount of adsorbed gas.

© 2007, first edition, EPFL Press

Corrosion and inhibition phenomena often involve the adsorption of reaction intermediates or dissolved species. The Langmuir isotherm generally serves as a starting point in the modeling of these phenomena as will be discussed in later chapters of this book.

Adsorption rate

The kinetic theory of gases allows us to predict the flux N_i of gas molecules that enter into collision with the surface.

$$N_i = \frac{P_i}{\left(2\pi M_i RT\right)^{1/2}} \quad (\mathrm{mol\,m^{-2}\,s^{-1}}) \tag{3.44}$$

M_i designates the molar mass of the gas and P_i its partial pressure. Only a fraction s_K of the molecules that collide with the surface can transfer their kinetic energy to the atoms of the solid and, by this process, remain adsorbed. We call s_K the ***sticking coefficient***. The adsorption rate is thus equal to

$$v_{ads} = s_K\, N_i = s_K \cdot \frac{P_i}{\left(2\pi M_i RT\right)^{1/2}} \quad (\mathrm{mol\,m^{-2}\,s^{-1}}) \tag{3.45}$$

with $0 < s_K < 1$.

It is instructive to estimate the time τ_m required for a monolayer of gas to become adsorbed onto a metal surface. To simplify, we assume that most of the molecules that hit the surface stick, i.e., $s_K \simeq 1$. A surface consists on average of 10^{19} adsorption sites/m^2. The flux of particles that strikes the surface is calculated by multiplying (3.44) by Avogadro's number N_L. This gives, at 300 K, for O_2 (M = 32 g mol^{-1}) a flux of $2.7 \times 10^{22}\ P_i$ [molecule m^{-2} s^{-1}]. The coverage time τ_m corresponds to the number of sites divided by the number of collisions. Thus, for a partial pressure of $P_{O2} = 10^5$ Pa (1 atm): $\tau_m \approx 10^{-9}$ s. At lower partial pressures, for example for $P_{O2} = 10^{-4}$ Pa (10^{-6} torr): $\tau_m \approx 1$ s. This result shows that even by working at reduced pressure, it is difficult to prevent a metal surface from becoming covered by an adsorbed gas characterized by a high sticking coefficient. It is not surprising, therefore, that in practice adsorbed species and/or thin oxide films almost always cover the surface of metals.

With the kinetic theory of gases, we can relate the adsorption coefficient to the sticking coefficient. Equations (3.35) and (3.45) give, in fact, for a pure, unexposed surface ($\theta_i = 0$):

$$k_{ads} = \frac{s_k}{\left(2\pi M_i RT\right)^{1/2}} \tag{3.46}$$

According to Figure 3.9, if the activation energy is zero for adsorption, it is equal to Q_{ads} for desorption. The adsorbed particle cannot escape the surface unless its thermal energy exceeds this value. By analogy to Arrhenius' Law, it is possible to create a model for the desorption rate constant k_{des}:

© 2007, first edition, EPFL Press

$$k_{\text{des}} = k_{\text{h}} \exp\left(-\frac{Q_{\text{ads}}}{RT}\right) \tag{3.47}$$

The factor k_{h} replaces the expression kT/h, with h, the Planck constant, and k, the Boltzmann constant. The definition $b = k_{\text{ads}}/k_{\text{des}}$ combined with equations (3.42) and (3.43) gives

$$b_{\text{L}} = \frac{s_{\text{K}} \cdot \exp\left(-\frac{Q_{\text{ads}}}{RT}\right)}{k_{\text{h}}(2\pi M_{\text{i}} RT)^{1/2}} \tag{3.48}$$

Thus an increase in temperature decreases the value of b_{L} and therefore favors the desorption of adsorbed species.

Experimental determination of the heat of adsorption.

The study of isotherms allows us to experimentally determine the value of Q_{ads}. The formula (3.49) represents the adsorption of a gaseous species i, and the expression (3.50) describes the variation of free energy associated with the phenomenon.

$$B_{\text{i(g)}} \rightleftharpoons B_{\text{i(ads)}} \tag{3.49}$$

$$\text{d}G = V^{\text{m}}_{\text{(ads)}}\text{d}P - S^{\text{m}}_{\text{(ads)}}\text{d}T + \left(\frac{G}{\theta}\right)\text{d}\theta - V^{\text{m}}_{\text{(g)}}\text{d}P + S^{\text{m}}_{\text{(g)}}\text{d}T \tag{3.50}$$

V^{m} represents the molar volume, S^{m} the molar entropy of species i and θ the coverage of the surface (to simplify, the subscript i is omitted). At equilibrium, $\text{d}G = 0$.

$$(V^{\text{m}}_{\text{(ads)}} - V^{\text{m}}_{\text{(g)}})\text{d}P - (S^{\text{m}}_{\text{(ads)}} - S^{\text{m}}_{\text{(g)}})\text{d}T + \left(\frac{G}{\theta}\right)\text{d}\theta = 0 \tag{3.51}$$

A constant coverage implies that $\text{d}\theta = 0$. In addition, the molar volume of the adsorbed species is negligible with respect to that of the same species in the gas phase: $V^{\text{m}}_{\text{(ads)}} << V^{\text{m}}_{\text{(g)}}$. For a perfect gas: $V^{\text{m}}_{\text{(g)}} = RT/P$.

The entropy of adsorption is written as:

$$\Delta S_{\text{ads}} = S^{\text{m}}_{\text{(ads)}} - S^{\text{m}}_{\text{(g)}} = -\frac{Q_{\text{ads}}}{T} \tag{3.52}$$

By applying this last equation, along with the preceding hypotheses, to equation (3.51), we find:

$$-\frac{RT}{P}\text{d}P + \frac{Q_{\text{ads}}}{T}\text{d}T = 0 \tag{3.53}$$

A rearrangement yields the ***Clausius-Clapeyron equation***, valid when $\theta =$ constant:

© 2007, first edition, EPFL Press

$$\frac{\mathrm{d}\ln P}{\mathrm{d}T} = \frac{Q_{\mathrm{ads}}}{RT^2} \tag{3.54}$$

or in another form:

$$\frac{\mathrm{d}\ln P}{\mathrm{d}(1/T)} = -\frac{Q_{\mathrm{ads}}}{R} \tag{3.55}$$

With equation (3.55), Q_{ads} can be calculated from the adsorption isotherms measured at different temperatures. Figure 3.12 shows the adsorption isotherm of CO on a platinum surface with an orientation characterized by the Miller indices (111) [2].

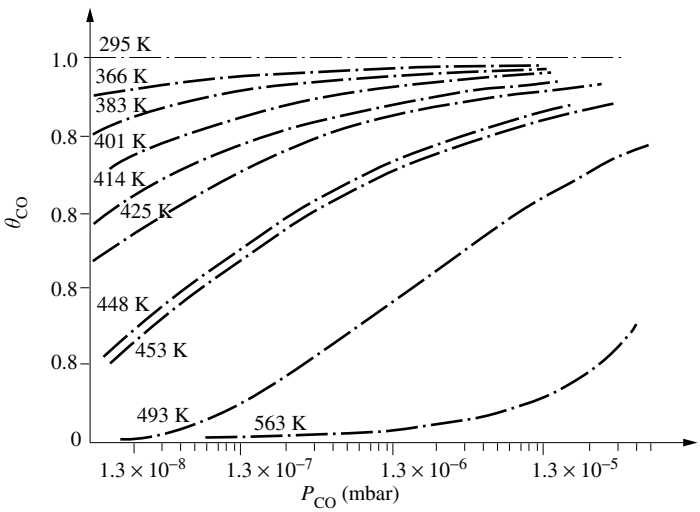


Figure 3.12 Adsorption isotherms of CO on Pt (111) (adapted from ref.[2]).

3.3 CHEMICAL CHARACTERIZATION OF SURFACES

3.3.1 Surface analysis methods

Adsorption, oxidation, and segregation phenomena can drastically modify the chemical composition of surfaces relative to that of the bulk and in this way influence the reactivity and the corrosion behavior of metals. Surface analysis methods refer to those experimental techniques that give chemical information on surfaces to a depth of 2–3 nm or less. Most surface analysis methods can be combined with ion sputtering, which progressively removes atomic layers from the surface. In this way surface analysis methods are able to measure concentration variations over a depth of tens or even hundreds of nanometers.

© 2007, first edition, EPFL Press

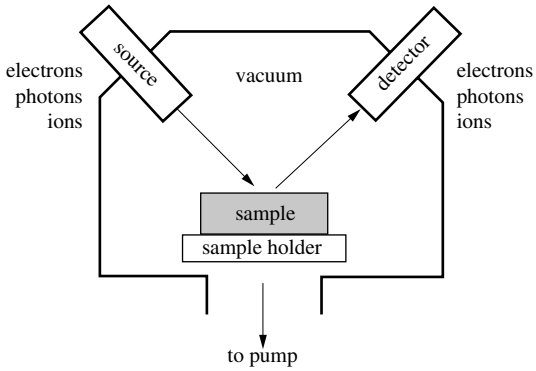


Figure 3.13 Principle of different surface analysis methods.

The principle of surface analysis methods is as follows (Figure 3.13). The surface to be analyzed is irradiated with electrons, X-rays or ions. As a result of collisions of the incident particles with atoms of the solid, electrons, photons or ions are emitted that contain specific chemical information. With a suitable detector one measures their quantity and their energy or mass.

Table 3.14 presents the four most common surface analysis methods, along with some of their characteristic features:

- X-ray photoelectron spectroscopy (XPS);
- Auger electron spectroscopy (AES);
- Secondary ion mass spectroscopy (SIMS);
- Ion scattering spectroscopy (ISS).

All of these techniques require an ultra-high vacuum chamber (10^{-6} to 10^{-8} Pa) to avoid contamination of the surface by undesired adsorption processes during the analysis. Vacuum conditions also limit the probability of accidental collisions between the emitted particles and gas molecules, which would lead to energy losses of the particles.

Table 3.14 Characteristics of common surface analysis methods

	AES	XPS	SIMS	ISS
Incident radiation (energy in KeV)	electrons (1–10)	photons (1–2)	ions (0.5–10)	ions (0.5–2)
Analyzed radiation	Auger electrons	photoelectrons	secondary ions	scattered ions
Analyzed depth (atomic layers)	2–10	3–15	2–10	1–2
Detection limit (% atomic layers)	0.1–1	0.1–2	10^{-4} –1	0.1–1
Quantitative analysis	yes	yes	limited	partially
Special characteristics	high spatial resolution, depth profiling	detection of oxidation state, non-destructive	high sensitivity, isotope detection	atomic layer resolution

© 2007, first edition, EPFL Press

A number of other physical methods are employed by corrosion scientists for the chemical characterization of surfaces and thin films: nuclear microanalysis (analysis of particles generated from nuclear reactions induced by a beam of high energy ions), Rutherford backscattering (elastic scattering of high energy ions), FTIR (Fourier-transform infrared spectroscopy), Raman spectroscopy, ellipsometry (an optical method based on the analysis of the polarization state of scattered light) or EXAFS (extended X-ray absorption fine structure analysis, a method using synchrotron radiation). Because these methods are of more specialized use in corrosion, we will not discuss them here.

3.3.2 Photoelectron spectroscopy

Photoelectron spectroscopy, commonly referred to as ***XPS*** (X-ray photoelectron spectroscopy) or ***ESCA*** (electron spectroscopy for chemical analysis), takes advantage of the photo effect discovered by Einstein (Figure 3.15). By absorbing a photon, an atom receives an energy $h\nu$. This energy is dissipated by the emission of an electron, the photoelectron. In XPS, incident photons typically have an energy of 1 to 2 keV. For example, magnesium, a frequently used source, emits radiation K_α of 1253.6 eV. Such a high energy of the incident radiation results in the emission of electrons from an inner shells of the atom. Figure 3.15 shows the ejection of a photoelectron from the K shell of an isolated atom.

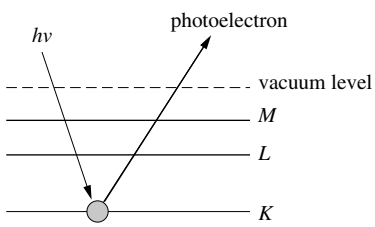


Figure 3.15 Principle of XPS (emission of a photoelectron).

Energy of XPS bands

In XPS one measures the kinetic energy of the photoelectrons emitted by the sample, E_{kin}. The electrical contact between the sample and the analyzer equalizes their electrical potential and, by this fact, their Fermi levels. It is therefore convenient to take the Fermi level as a reference point for the measurement. Thus, by arbitrarily setting $E_{\text{F}} = 0$, we can write an expression for the energy balance (Figure 3.16):

$$E_{\text{kin}} = h\nu - E_{\text{b}} - \Phi_{\text{w,det}} \tag{3.56}$$

Here E_{b} is the ***binding energy of the electron***, which corresponds to the energy difference between the Fermi level and the energy level from which the photoelectron is extracted. The binding energy of an electron in a given shell is characteristic of the

© 2007, first edition, EPFL Press

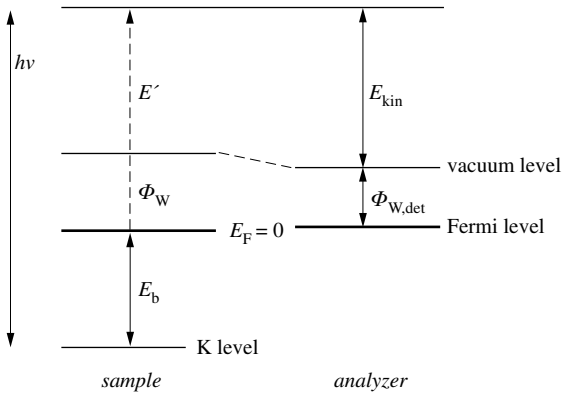


Figure 3.16 Energy of electrons in XPS measured with respect to the Fermi level.

chemical element that emits the photoelectron. For this reason XPS spectra are generally presented in terms of E_b rather than of E_{kin}.

The ***work function*** of the detector, $\phi_{W,det}$, characterizes the difference in energy between the Fermi level of the analyzer and vacuum. In practice, rather than determining this value, one prefers to calibrate the analyzer with a standard sample, such as gold, for which the photoelectron energies are known. XPS analysis is applicable to conducting and to non-conducting materials. However, if the conductivity of the sample is insufficient, the Fermi levels of the analyzer and the sample may differ, leading to a reduced precision of the energy bands. To avoid such effects the surface of insulator samples is sometimes flooded with low energy electrons.

Identification of XPS bands

Figure 3.17 shows the XPS spectrum of silver [3]. There are a number of peaks, because photoelectrons are emitted from several energy levels. The energy levels of

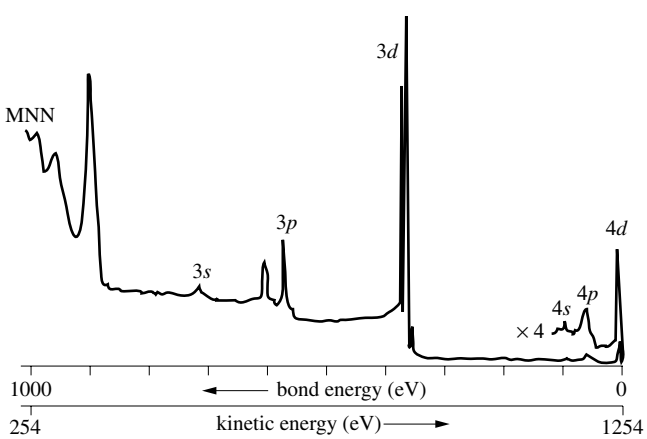


Figure 3.17 XPS survey spectrum of Ag; primary source is Mg K_α (adapted from ref. [3]).

© 2007, first edition, EPFL Press

electrons in an atom are defined by three quantum numbers: the principle quantum number $N_p = 1,2,3 \ldots$, which refers to the shells *K, L, M, . . .,* the secondary quantum number $N_l = 0, 1, 2, \ldots, (N-1)$, designating the orbitals *s, p, d, f,* and the spin quantum number $N_s = \pm 1/2$. In Figure 3.17, the double peak *3d* indicates that two electrons of opposite spin belonging to this state have been ejected. In the literature the designation of the energy states that correspond to a given set of quantum numbers is not uniform and depends on the field of application. Table 3.18 shows the notation generally employed for the designation of XPS bands and for X-ray emission, respectively. The quantum number $N_j = (N_l + N_s)$ shown in the table characterizes the total angular momentum of an electron.

Table 3.18 XPS bands and X-ray nomenclature.

N_p	N_l	N_j	XPS nomenclature	X-ray nomenclature
1	0	1/2	$1s_{1/2}$	K
2	0	1/2	$2s_{1/2}$	L_1
2	1	1/2	$2p_{1/2}$	L_2
2	1	3/2	$2p_{3/2}$	L_3
3	0	1/2	$3s_{1/2}$	M_1
3	1	1/2	$3p_{1/2}$	M_2
3	1	3/2	$3p_{5/2}$	M_3
3	2	3/2	$3d_{3/2}$	M_4
3	2	5/2	$3d_{5/2}$	M_5

Electron escape depth

A photoelectron that undergoes inelastic collisions loses a part of its energy. In this case, it is no longer possible to determine its quantum state of origin from the energy balance (3.56). The average distance that an electron travels in a solid without

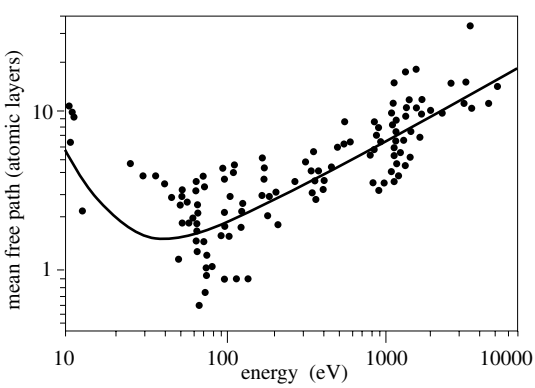


Figure 3.19 Variation of the mean free path of electrons in metals as a function of the kinetic energy (adapted from ref. [4]).

© 2007, first edition, EPFL Press

collision is the ***mean free path*** λ. Its value depends on the kinetic energy of the electron (Figure 3.19). In XPS, the mean free path of photoelectrons in the solid is on the order of one nanometer. The mean free path and the escape angle ϕ indicated in Figure 3.20 determine the ***escape depth*** Λ of the photoelectron:

$$\Lambda = \lambda \cdot \cos \phi \tag{3.57}$$

Photoelectrons that escape from depths less than Λ escape from the lattice without collision and thus contribute to XPS peaks. The other photoelectrons become part of the background noise.

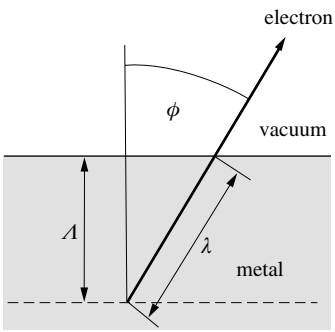


Figure 3.20 Electron escape depth Λ and mean free path λ of electrons.

Quantitative analysis

The following equation allows us to calculate the intensity I_i of a given XPS peak for the element *i*.

$$I_i = \kappa \int_0^\infty c_{i(z)} \exp\left(-\frac{z}{\Lambda}\right) dz \tag{3.58}$$

Here $c_{i(z)}$ is the concentration of the element *i* in the solid, and *z* the depth. The proportionality factor κ is obtained from the formula:

$$\kappa = \kappa_{sp} \cdot I_p \cdot \sigma_e \tag{3.59}$$

where κ_{sp} is a parameter dependent on the spectrometer used and its mode of operation; I_p represents the intensity of the incident radiation; and σ_ε is the ionization cross section or, expressed otherwise, the probability that an irradiated atom looses a photoelectron. To integrate (3.58), one needs first to decide which function best represents the concentration profile $c_{i(z)}$ near the surface. Figure 3.21 shows three examples: (a) constant concentration, (b) exponentially decreasing concentration, and (c) a concentration step function.

In the case of constant concentration (a), the integration of (3.58) gives:

$$I_i = \kappa \cdot \Lambda \cdot c_{i,b} \tag{3.60}$$

where $c_{i,b}$ corresponds to the concentration of the element *i* in the bulk.

© 2007, first edition, EPFL Press

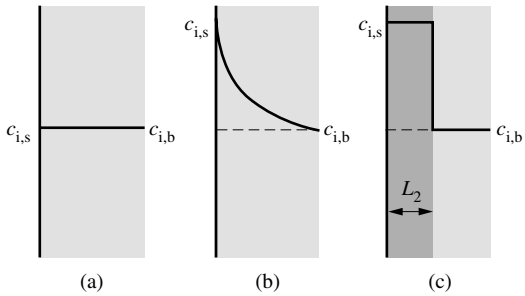


Figure 3.21 Typical concentration profiles near metal surfaces for the interpretation of XPS data: (a) uniform concentration; (b) exponential decay; (c) concentration step function.

The concentration profile shown in Figure 3.21(c) is frequently used to characterize thin oxide films of known thickness, L_2. The concentrations (assumed constant) of the element i are given by $c_{i,ox}$ in the oxide and by $c_{i,m}$ in the metal substrate. The integration of (3.58) yields the intensity of the signal emitted by the element i of the oxide and of the metal, respectively:

$$I_{i,ox} = \kappa \Lambda c_{i,ox} \left[1 - \exp\left(-\frac{L_2}{\Lambda} \right) \right] \tag{3.61}$$

$$I_{i,m} = \kappa \Lambda c_{i,m} \exp\left(-\frac{L_2}{\Lambda} \right) \tag{3.62}$$

Depending on the phase to which it belongs (i.e., metal or oxide), a given element generally emits photoelectrons of slightly different energy. Thus we can distinguish between the intensities $I_{i,m}$ and $I_{i,ox}$. If κ and Λ are known, we obtain a system of two equations, (3.61) and (3.62), and three unknowns, $c_{i,ox}$, $c_{i,m}$ and L_2. Provided the concentration of the substrate $c_{i,m}$ underneath the film is known, one can determine the concentration in the oxide film and its thickness. The XPS method therefore is particularly well suited for the study of thin passive oxide films on metals.

Chemical shift

Even though electrons from inner orbitals do not participate in chemical bonds, their energy state is dependent to some extent on the nature of the bonding. The kinetic energy of photoelectrons emitted from inner levels, therefore, varies slightly with the binding state of the emitting atom. This phenomenon, called the ***chemical shift,*** is easily measured by comparing the XPS spectra of the metallic and oxidized states of a specific element.

Figure 3.22 shows the peak $3p_{3/2}$ of chromium measured on an iron-chromium-molybdenum alloy surface covered by thin oxide film [5]. The background noise has been subtracted, and the double peak of chromium has been deconvoluted into two peaks that correspond to oxidized and metallic chromium. Even though this procedure

© 2007, first edition, EPFL Press

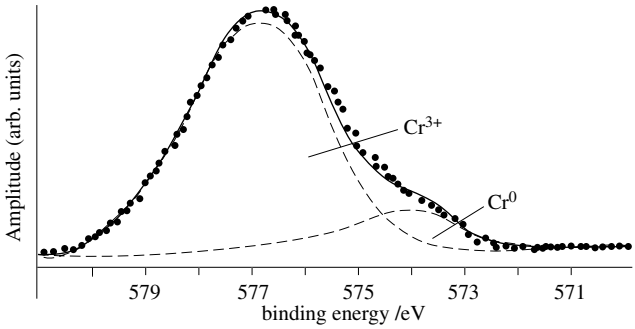


Figure 3.22 Deconvolution of XPS peak Cr $3p_{3/2}$ as measured on a passive film that formed on an iron-chromium-molybdenum alloy (adapted from ref. [5]).

requires a number of assumptions, in particular concerning the shape of the bands, it allows us to determine an intensity for each of the two bands. Table 2.23 gives the values for the electron binding energies of the elements Fe, Cr and Mo in their metallic and oxidized states observed on a Fe-Cr-Mo alloy with a thin passive oxide film. A chemical shift on the order of 3–5 eV is observed.

Table 3.23 Electron binding energies in surface films and in metals measured by XPS on a passive Fe-Cr-Mo alloy (adapted from ref. [6]).

Peak	Metal (eV)	Oxide (eV)	Oxidation state	Chemical shift (eV)
Fe $2p_{3/2}$	707.1	710.7	3+	3.6
Cr $2p_{3/2}$	574.2	577.1	3+	2.9
Mo $3s_{5/2}$	227.8	232.5	6+	4.7

As a general rule, the binding energy of core level electrons increases with the degree of oxidation. In relative terms, however, chemical shifts are weak. The values shown in Table 3.23 do not exceed 2% of the binding energy.

3.3.3 Auger Electron Spectroscopy

In Auger electron spectroscopy, abbreviated ***AES***, one determines the number and the energy of Auger electrons. The principle of AES is the following. Irradiation of a solid with X-rays or with an electron beam of suitable energy leads to the ejection of an electron from the inner shell of the target atom. The core level vacancy thus created is immediately filled by an electron from an outer shell. This liberates energy, which can be dissipated either by the emission of an X-ray photon or by emission of an ***Auger electron.*** In the latter case, the excess energy is transmitted to an outer shell electron, which thus acquires enough kinetic energy to be ejected from the atom. The two energy dissipation processes are schematically shown in Figure 3.24. The probability to observe one or the other depends primarily on the atomic number of the irradiated atom. Elements with small atomic number are more likely to emit an Auger

© 2007, first edition, EPFL Press

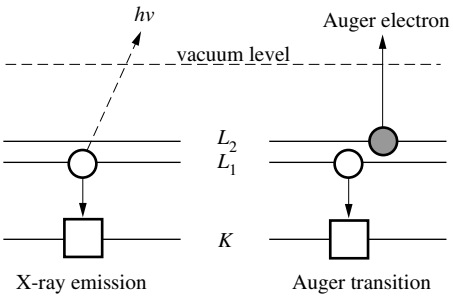


Figure 3.24 The emission of an X-ray or of an Auger electron.

electron, elements of high atomic number an X-ray photon (Figure 3.25). Auger electron spectroscopy, therefore, is particularly sensitive to light elements, while methods that take advantage of X-ray emission, such as the *electron microprobe* are more sensitive to heavy elements. the electron microprobe is commonly used for microanalysis of metallic samples, but because the analyzed X-rays originate from a depth of some micrometer it is not a true surface analysis method.

Auger transitions, i.e., electronic transitions followed by the emission of an Auger electron are described according to the quantum states of the three electrons involved using the X-ray nomenclature (Table 3.18). In the transition KL_1L_2 shown in Figure 3.24 an electron from level L_1 has filled a hole in level K, resulting in the emission of an Auger electron from the level L_2. Often, the subscripts are dropped: instead of KL_1L_2, we simply write KLL.

Figure 3.25 refers to the ionization of the K level of an atom, but Auger transitions of the type LMN and MNN are also common. They result from ionizations of the L and M energy levels, respectively and they take place with a high probability in elements of large atomic number. As a consequence, Auger electron spectroscopy is not limited to light elements, but is also well suited for the analysis of heavier elements of the periodic table.

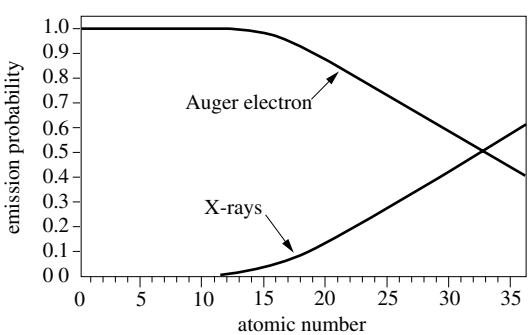


Figure 3.25 Probabilities of X-ray or Auger-electron emission during the K ionization of different elements (adapted from ref. [7]).

© 2007, first edition, EPFL Press

Energy of Auger peaks

The kinetic energy of Auger electrons is obtained in a similar way as already seen for photoelectrons. Acting as a reference point, the Fermi level is arbitrarily assigned the value of zero. The energy balance for any given transition *WYZ* (Figure 3.26) yields:

$$E_W - E_X \approx E_Y + E_{kin} + \Phi_{W,det} \tag{3.63}$$

where E_W, E_X and E_Y are the electron binding energies at the levels *W*, *X* and *Y* respectively, E_{kin} represents the kinetic energy of an Auger electron, and $\phi_{W,det}$ is the work function of the analyzer.

In reality, the energy E_Y does not correspond exactly to the electron binding energy in the level *Y* of the unperturbed atom. Its value is somewhat greater, because the missing negative charge resulting from the creation of electronic holes slightly increases the force exerted by the nucleus on the electrons in level Y. To take account of this effect, the correction term Δ is introduced into equation (3.63). We thus obtain the kinetic energy $E_{kin} = E_{WXY}$ for the Auger electron produced by the transition *WXY*:

$$E_{WXY} = E_W - E_X - E_Y - (\Phi_{W,det} + \Delta) \tag{3.64}$$

The value of the term in parentheses is small relative to E_{WXY}. In practice, measured energies are calibrated against a standard sample; as a result, it is not necessary to know the value of $(\phi_{W,det} + \Delta)$.

In contrast to photoelectrons, the kinetic energy E_{WXY} of Auger electrons does not vary with the energy of the incident radiation. For this reason, in Auger spectra we normally indicate the kinetic energy of the analyzed electrons, rather than the binding energy generally used in XPS. Auger electrons and photoelectrons have kinetic energies of the same order of magnitude, and their mean free paths are comparable. The significance of this observation is that analyses using one or the other of the two methods will probe surface layers of similar thickness.

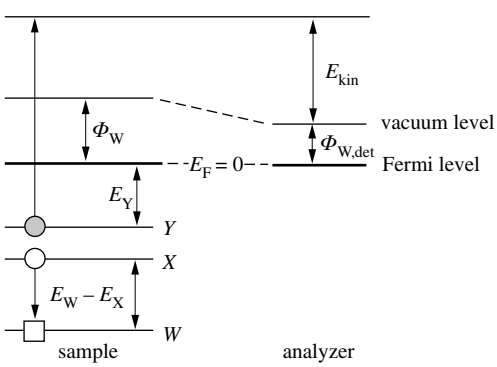


Figure 3.26 Auger transition WXY.

© 2007, first edition, EPFL Press

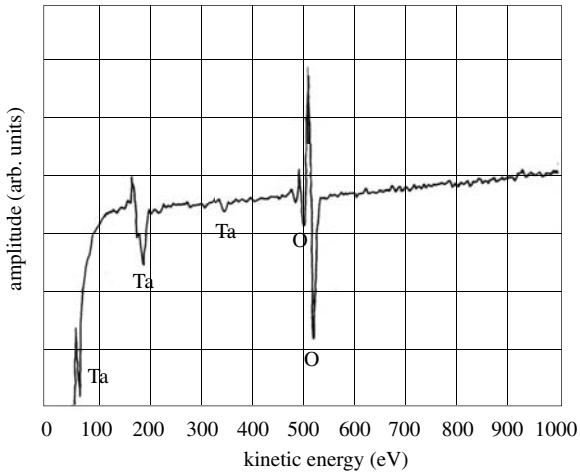


Figure 3.27 AES spectrum of tantalum oxide taken in the differential mode (adapted from ref. [8]).

Auger electrons can be analyzed with the same detection techniques used in XPS. One thus obtaines spectra showing the number of Auger electrons as a function of their kinetic energy, N(E). Because Auger transitions yield small signals in a steep background noise, Auger spectra are often presented in the derivative form: dN(E)/dE. Certain analyzer types (cylindrical mirror analyzer) yield directly differential spectra. Figure 3.27 shows the differential Auger spectrum of a tantalum surface covered with a thin oxide film.

Quantitative analysis

Quantitative analysis using Auger electron spectroscopy is based on equation (3.58), developed originally for XPS. An important modification is the substitution of the proportionality constant κ by κ':

$$\kappa' = \kappa_s \, I_p (1+r) \sigma_e \, \gamma_{AES} \tag{3.65}$$

Here, κ_s is a instrument-specific calibration factor, I_p is the intensity of the incident electron beam, r represents the scattering factor for electrons in the solid, σ_e is the ionization cross section, and γ_{AES} is the probability of an Auger transition.

In practice one often uses ***relative elemental sensitivity factors*** for quantifying AES or XPS data. One assumes that the intensity of a given AES or XPS signal is proportional to the concentration of the element in the analyzed sample multiplied by its relative elemental sensitivity factor. The latter is an experimental quantity applying to specific operating conditions. The analysis software of most commercial AES and XPS instruments includes databases of sensitivity factors that allow one to compute the elemental concentration.

© 2007, first edition, EPFL Press

Depth profile analysis

AES and XPS both probe surfaces to a depth in the nanometer range. And yet, in order to understand certain corrosion mechanisms, we often would like to know the composition of much deeper layers and observe the variation in composition as a function of depth. To achieve this goal one combines surface analysis with ion sputtering. Bombardment of a surface with ions, most often Ar^+ or other noble gas ions, in the energy range between 1 et 10 KeV leads to progressive removal of surface layers. Continuously or at selected time intervals, one monitors the intensity of the Auger or XPS signal of the elements of interest (Figure 3.28). The resulting intensity–sputter time data can then be converted into a concentration-depth profile using relative elemental sensitivity factors providing one knows the sputter rate. For pure metals sputter rates are easily obtained from the crater depth measured after a given time of sputtering. However, in presence of oxide films or concentration gradients the exact sputter rate is often not known. Figure 3.29 shows the Auger depth profile of a tantalum oxide film, 30 nm thick, prepared by the anodization of tantalum. The oxide-metal interface and the presence of a carbon contamination on the surface are clearly seen. In a similar fashion, XPS analysis combined with ion sputtering yields XPS depth profiles.

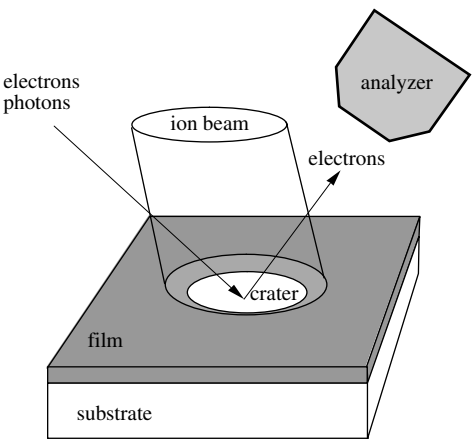


Figure 3.28 Principle of AES or XPS depth profile analysis.

Chemical Imaging

AES as well as XPS can be used for surface imaging, but the spatial resolution of AES is generally better. The reason is that the primary electron beam used for exciting the surface atoms in AES can be well focused on areas as small as a few nanometer. X-rays are more difficult to focus and additional techniques, not discussed here, are required to improve the spatial resolution in XPS, which for most instruments today is in the ten micrometer range. In order to acquire a chemical image an electronic guiding system scans the surface while the AES or XPS intensity of chosen peaks is measured and stored numerically. The chemical map of the surface thus produced

© 2007, first edition, EPFL Press

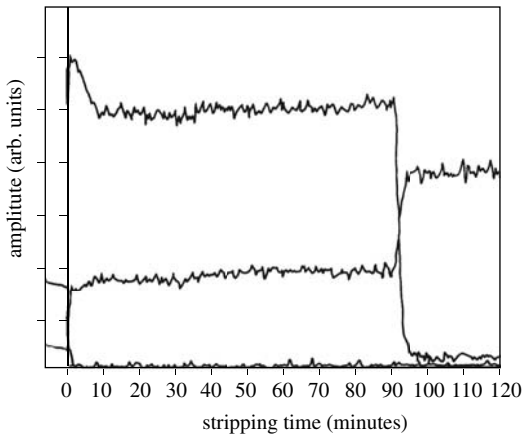


Figure 3.29 AES depth profile of a 30-nm-thick tantalum oxide film on tantalum, prepared by anodization (adapted from ref. [8]).

represents the spatial distribution of these elements on the surface. For elements exhibiting well pronounced chemical shifts the spatial distribution of different oxidation states can also be imaged. AES and XPS chemical images differ from those obtained with an electron microprobe by the fact that the information comes from a much smaller depth.

3.3.4 Ion scattering spectroscopy

In ion scattering spectroscopy, commonly referred to as ***ISS***, one obtains information about a surface by measuring the energy and quantity of elastically scattered ions. When an ion beam having an energy on the order of 1 keV irradiates a surface, both elastic and inelastic collisions take place between the incident particles and the target atoms. The speed of the elastically scattered ions, i.e., the ions that bounce back, depends on their energy before the collision and on the mass of the atoms undergoing collision. By setting up expressions for energy and momentum balances, one can calculate the mass, and thus the identity, of the irradiated atoms.

Figure 3.30 schematically shows the process of scattering. An incident ion of mass m_1 and of kinetic energy E_p, undergoes an elastic collision with an atom of mass m_2 located at the surface of the solid.

$$E_p = \frac{1}{2} m_1 v_p^2 \tag{3.66}$$

with v_p representing the velocity of the incident ion.

After the collision, the scattered ion has velocity v_1 and energy E_1. As for the collided atom, it has acquired velocity v_2. The conservation of energy allows us to set up the following balance:

© 2007, first edition, EPFL Press

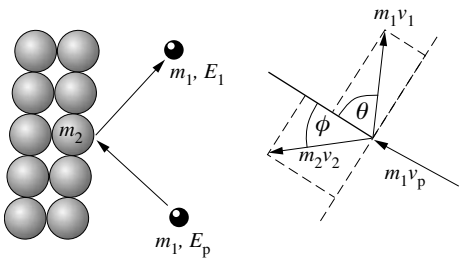


Figure 3.30 Elastic collision between a gas ion of mass m_1 with a surface atom of mass m_2.

$$\frac{1}{2}m_1v_p^2 = \frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2 \tag{3.67}$$

The conservation of momentum parallel and perpendicular to the direction of the incident beam yields:

$$m_1v_p = m_1v_1 \cos\theta + m_2v_2 \cos\phi \tag{3.68}$$

$$0 = m_1v_1 \sin\theta + m_2v_2 \sin\phi \tag{3.69}$$

By eliminating the unknowns v_2 and ϕ, we can calculate the energy ratio E_1/E_p:

$$\frac{E_1}{E_p} = \left[\frac{m_1 \cos\theta + (m_2^2 - m_1^2 \sin^2\theta)^{1/2}}{m_1 + m_2}\right]^2 \tag{3.70}$$

In practice, in order to simplify (3.70), $\theta = 90°$ is often selected:

$$\frac{E_1}{E_p} = \frac{m_2 - m_1}{m_2 + m_1} \tag{3.71}$$

Equation (3.71) shows that the ratio E_1/E_p depends on the difference in mass between m_2 and m_1. With the known mass m_1, one needs to measure E_1 in order to determine m_2. The smaller m_1, the more precise the measurement will be, explaining the choice of He^+ as the primary ion.

In order to distinguish between two neighboring elements of the periodic table, the energy of the scattered ions must be known precisely. According to (3.71), the required precision is all the higher as the masses of the involved particles is high. Therefore, the ISS method is most easily applied to elements of low atomic mass.

The flux of scattered ions of energy E_1 depends directly on the concentration of the element of mass m_2 at the surface. It is therefore possible to interpret the results of ISS analyses in a semi-quantitative, or even quantitative, fashion.

The ISS method is the only surface analysis method that is strictly concerned with the first atomic layer of a material. Nevertheless, under certain conditions, the primary

© 2007, first edition, EPFL Press

beam can induce sputtering of the surface, allowing a concentration-depth profile to be measured. Because of some disadvantages such as a limited precision for atomic masses and concentrations and a high detection threshold, ISS has found fewer applications in the study of corrosion problems than either XPS or AES.

3.3.5 Secondary ion mass spectroscopy

In secondary ion mass spectroscopy, abbreviated ***SIMS***, one monitors the ionic species that are emitted from a surface as a result of irradiation with primary ions in the energy range of 0.5–10 KeV. When a primary ion in this energy range strikes the surface of a solid, one of two phenomena can take place, depending on its energy. An elastic collision may transform the primary ion into a scattered ion as described above for ISS. On the other hand, the primary ion may penetrate into the solid (about 1–2 nanometers) whereby it undergoes a number of elastic and inelastic collisions (Figure 3.31). It thus progressively looses its kinetic energy, which is transferred to the atoms of the solid situated in a zone close to the surface. The energy transfer sets off a chain of atomic collisions there, which modifies the crystal lattice in the surface zone, even to the point of converting it into an amorphous structure. The described phenomenon is referred to as *atomic mixing*. As a consequence of these events, a number of surface atoms acquire enough kinetic energy to be ejected from the solid. A small portion of the ejected species (about 1% or less) carries a positive or negative charge, these are called ***secondary ions***. With the help of a mass spectrometer one determines the mass/charge ratio (m_i/q_i) of the secondary ions and their quantity. A typical SIMS equipment comprises an ultra high vacuum chamber equipped with a primary ion source, and a mass spectrometer.

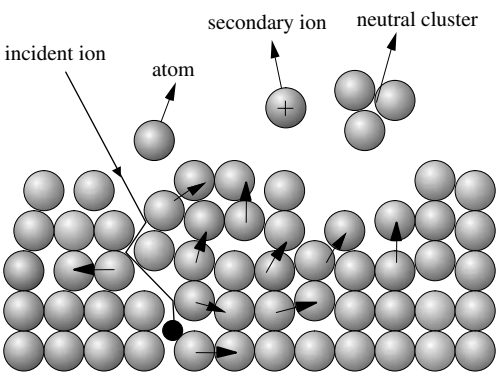


Figure 3.31 Elastic and inelastic collisions leading to emission of secondary ions

Static and dynamic modes

In SIMS, we distinguish between static and dynamic modes. In ***static SIMS*** one uses a primary beam with a very low current density (≤ 10 nA cm^{-2}), which leaves the surface largely unperturbed. In order to obtain sufficient signal intensity, a

© 2007, first edition, EPFL Press

relatively large-diameter beam, on the order of millimeters, is used. The weak intensity of the incident beam drastically reduces the sputtering rate, resulting in negligible material loss of the sample during analysis. Static SIMS is a largely non-destructive technique, suitable for the study of adsorption phenomena and for chemical surface analysis. Figure 3.32 shows the spectrum of an iron surface containing adsorbed CO [8]. It is easy to recognize the different ionized species of iron, carbon, and oxygen. The results suggest that CO adsorbs onto iron in molecular and in dissociated form.

Static SIMS analysis provides much information about the molecular structure of adsorbed species and the chemical composition of surfaces. However, the interpretation of SIMS spectra is often not straight forward. The reason is that the spectra include ionic species that are the result of gas phase collision reactions taking place in the immediate vicinity of the surface.

SIMS in the dynamic mode, ***dynamic SIMS***, is particularly useful for the measurement of concentration-depth profiles. A primary beam of small diameter ($\leq$ 1–10 μm) sweeps the surface. Its intensity, much greater than that applied in static SIMS, generally exceeds 10 μAcm^{-2}, sufficient to sputter the surface. To avoid crater edge effects one uses electronic gating, which permits to analyze only in the center of a sputter crater. Most often O_2^+ or Cs^+ ions are employed for the primary beam because they provide a high secondary ion yield. An O_2^+ ion source results in the production of positive secondary ions, a Cs^+ ion source of negative ones. Dynamic SIMS is also well suited for chemical imaging of surfaces (SIMS microanalysis), by use of an integrated scanning system for the primary ion beam.

The main difference between static and dynamic SIMS lies in the ion dose that the sample receives during a measurement. In static SIMS it typically lies below $< 10^{12}$ ions cm^{-2}, which means that there is about one incident ion per 1000 surface atoms. In this way one considers the perturbation of the surface by the ion beam negligible. Instead of reducing the primary current density one can also limit the ion dose by reducing the time of the perturbation. This method is used in time-of-flight SIMS, abbreviated ***TOF-SIMS***. The primary beam is pulsed and the secondary ions are separated according to their time of flight between the sample and the detector.

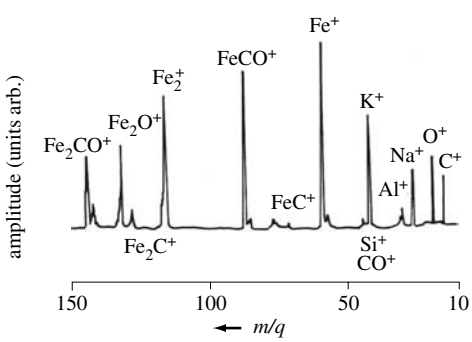


Figure 3.32 Mass spectrum of an iron surface containing absorbed CO (adapted from ref. [9]).

© 2007, first edition, EPFL Press

TOF-SIMS not only allows one to perform static SIMS experiments, but many instruments also permit chemical imaging of surfaces as well as depth profile analysis.

Quantitative analysis

Starting with equation (3.72), it is possible, under certain conditions, to quantitatively interpret the results obtained with SIMS.

$$I_i = I_p\, Y_i\, \alpha_i\, c_i\, T_{sp} \tag{3.72}$$

I_i represents the intensity of the signal m_i/q_i; I_p refers to the primary beam intensity; c_i is the atomic concentration of the element *i*; and T_{sp} is a proportionality factor whose value depends on the equipment used and the operating conditions.

The sputter yield Y_i expresses the ratio ejected atoms / incident ions. Its value may vary typically between 0.5 and 10. A sputter yield greater than 1.0, means that a single incident ion is capable of ejecting several surface atoms by the previously discussed mechanism of chain collisions (Figure 3.31).

The ionization probability α_i represents the fraction of ejected atoms, belonging to species *i,* that are ionized. Depending on the chemical nature of the solid and the operating conditions, α_i can vary between 10^{-1} and 10^{-5}. The ionization probability strongly depends on the chemical nature of the surface. This greatly complicates, and sometimes makes impossible, a quantitative interpretation of SIMS results. The problem is particularly pronounced when analysing thin oxide films on alloys. On the other hand, SIMS is an excellent method for measuring concentration profiles of trace elements in matrices of almost uniform composition, and therefore the method is widely applied to the analysis of doping elements in semiconductors.

Despite the inconveniences resulting from the strong matrix effects, the SIMS method finds numerous applications in corrosion. Because of its high sensitivity (at the ppm level), SIMS is particularly suited for the study of species present in very small concentration, such as trace analysis in metals and thin oxide films. SIMS can be applied to conducting and non conducting solids, including organic coatings, where SIMS can yield specific chemical information about surface or interface species present. Another interesting application are isotope studies. The high mass resolution of most mass spectrometers used in SIMS permits to monitor the quantity of different isotopes of a given element. This opens the way for SIMS to be used for marker studies with isotopes, for example, for the investigation of the growth mechanism of oxide films.

3.4 STRUCTURAL CHARACTERIZATION OF SURFACES

3.4.1 Atomic structure and topography of metal surfaces

When discussing the structure and topography of metal surface we shall distinguish two basic structural scales:

© 2007, first edition, EPFL Press

- the microscopic scale;
- the atomic or nanometer scale.

Surface structure and topography at the microscopic scale

The topography of metal surfaces at the microscale depends above all on the manufacturing process: machining, grinding polishing, rolling, chemical etching, etc. For example, mechanical machining with cutting tools results in strong plastic deformation of the metal close to the surface and usually leads to a rugosity of many micrometers, depending on the machining parameters and the properties of the material. Mechanical polishing results in a much smaller rugosity, well below a micrometer, but it is also causes some a plastic deformation. The depth of the deformed zone (the so-called Beilby layer) is about 10 times that of the final surface roughness. Chemical or electrochemical polishing provide a means to produce smooth surfaces without any plastic deformation, but under unfavorable conditions these processes cause the formation of pits or shallow surface waviness referred to as "orange peel" topography.

Non-metallic inclusions or second phase precipitates also influence the topography of surfaces. Some types of chemical attack highlight inhomogeneities by dissolving preferentially the metal, while others lead to preferential dissolution of the inclusions, and to pit formation. On the other hand, thin passive oxide films of nanometer thickness, have no effect on the topography of metal surfaces at the microscopic scale. Generally, smooth surfaces on the microscopic scale are more resistant to corrosion than rough surfaces because roughness facilitates the formation of microscopic corrosion cells and the local breakdown of thin oxide films (Chapter 7). Most of all, corrosion products adhere less well on smooth surfaces.

Surface structure and topography at the atomic scale

At the atomic scale, the structure of metal surfaces depends on the crystal orientation and on the presence of crystalline defects. We shall distinguish three kinds of surfaces:

- densely packed low-index surfaces;
- slightly misoriented surfaces;
- complex surfaces.

Densely packed surfaces correspond to low Miller indices, as shown in Figure 3.1 for an fcc crystal structure. The drawing shows that the [111] surface plane is most densely packed. The minimization of surface tension, discussed at the beginning of this chapter, generally favors formation of densely packed surfaces. On arbitrary oriented crystals this behavior produces faceting when the metal is heated in absence of surface contamination. Similarly, in the absence of adsorption effects the electrochemical dissolution or deposition of metals tends to expose low index surfaces.

Slightly misoriented surfaces are oriented at a small angle with respect to a low index plane. They are generally described by the ***TLK*** model, a name that comes from

© 2007, first edition, EPFL Press

the words *terrace, ledge, kink*. Figure 3.33 shows a slightly misoriented surface that consists of a sequence of dense planes, the terraces, which are separated by monoatomic steps or ledges. The distance between the ledges determines the angle β, which characterizes the macroscopic orientation of the surface with respect to the compact plane. As for the kinks, they induce a slope in a second direction.

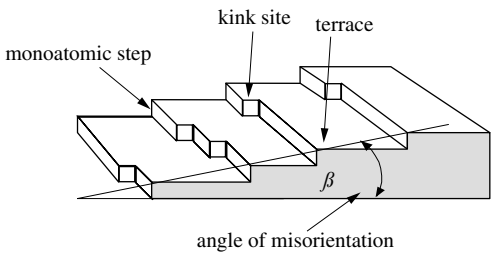


Figure 3.33 Terrace-Ledge-Kink (TLK) model of metal surfaces.

Complex surfaces, are characterized by their strong misorientation with respect to compact planes. As a consequence the atomic steps are very closely spaced and the TLK model is not readily applied. The absence of well-defined geometric sequences prevents a simple description of complex surfaces.

Most metals and alloys used by engineers are polycrystals. In addition to dislocations and point defects, their surfaces are crisscrossed by grain boundaries that separate regions of different orientation and, in case of alloys, by phase boundaries that separate different metallic phases. The atoms located at these defects are less solidly bound in the crystal lattice, and therefore they are favorite sites for chemical or electrochemical attack. The phenomenon is used in metallography to visualize the microstructure of metals and alloys. Figure 3.34 shows a micrograph of a polycrystalline copper surface that has been subjected to electrochemical dissolution. The grain boundaries are clearly visible.

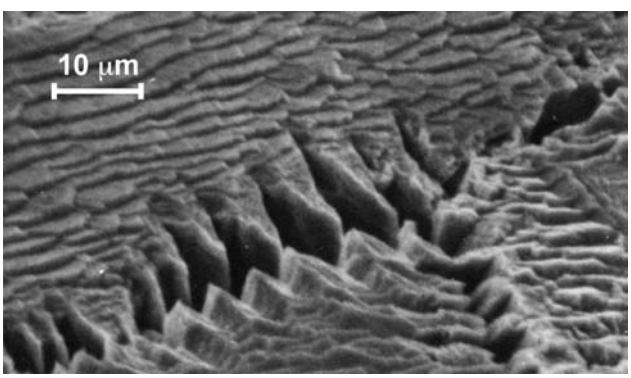


Figure 3.34 Scanning electron micrograph of polycrystalline copper surface after anodic attack.

© 2007, first edition, EPFL Press

3.4.2 Methods for the characterization of surface topography

A number of different methods are used for the characterization of surface structure and topography. Table 3.35 shows methods that provide information at the atomic scale, the nanometer scale and the micrometer scale, respectively. In practice, however, the distinction is not so clear; the scanning electron microscope (SEM) is also well suited for the study of micrometer scale or even millimeter scale surface topography, and the atomic force microscope (AFM) can achieve atomic resolution under favorable circumstances. Furthermore, when discussing resolution one should distinguish between the vertical and the lateral resolution, which may differ greatly for a given method. For example, the lateral resolution of local probe instruments such as the AFM is limited by the dimension of the tip that limits its capability to penetrate narrow recesses on a rough surface and that prevents it from accurately reproducing a steep wall. Similarly, the interference microscope has submicrometer depth resolution, but can be used only on polished surfaces because strong roughness renders the observation of interference fringes impossible.

Table 3.35 Methods for structural characterization of surfaces

Observation	Method
atomic scale	low energy electron diffraction (LEED)
	field ion microscopy (FIM)
	scanning tunnel microscopy (STM)
	high resolution transmission electron microscopy
submicrometer scale	atomic force microscopy (AFM)
	scanning electron microscopy (SEM)
	interference microscopy
micrometer to millimeter scales	scanning electron microscopy (SEM)
	stylus methods (diamond tip, laser probe)
	optical microscopy

Structural characterization of surfaces at the atomic scale

Low energy electron diffraction (LEED) is used to study the crystalline structure of pure metal surfaces or of metal surfaces covered with adsorbed species. The method is based on the interpretation of diffraction patterns created by elastically scattered electrons. The primary electrons irradiating the sample have an energy that varies between 10–500 eV, which corresponds to a wavelength of 0.4 to 0.05 nm. Because of this, the electrons barely penetrate into the crystal, typically only to distances of 1–2 nm. The diffraction patterns therefore provide information about the arrangement of the first 5 to 10 atomic layers.

The LEED method requires ultra high vacuum to avoid collisions between the scattered electrons and residual gas atoms and to prevent contamination of the sample surface during a measurement. Essentially limited to the study of monocrystalline surfaces, LEED is little used for corrosion studies.

© 2007, first edition, EPFL Press

Field ion microscopy, or ***FIM***, is a classical technique for the study of surface structure that can yield a structural image with a magnification factor on the order of 10^7. About 10,000 V are applied between the sample, shaped into a pointed tip (positive pole), and a fluorescent screen in a low-pressure atmosphere of helium. The helium atoms that approach the sample tip loose an electron due to the strong electric field and the resulting He^+ ions are accelerated towards the fluorescent screen where they form an image. The ionization probability depends on the local intensity of the electric field at the sample, and therefore on the atomic structure. The method is limited to refractory metals such as tungsten, tantalum or iridium.

The ***scanning tunneling microscope*** or ***STM*** invented in 1982 by Binnig and Rohrer [9] has since revolutionized the study of surfaces at the atomic and the nanometer scales. Quantum mechanics shows that the electrons in a metal can tunnel through the potential barrier situated at the surface. The probability of finding an electron outside the solid decreases exponentially with distance and drops to essentially zero at about 1 nm. By applying an electric voltage between the tip (ideally having a point consisting of a single atom) and the sample surface, a tunneling current is measured whose intensity decreases exponentially with the distance between the two objects.

$$I \propto \frac{U}{d} \exp\left(-k\ \Phi_W\ d\right) \tag{3.73}$$

Here, U is the applied voltage, d is the separation distance, k is a constant ($\approx$ 10.25 $eV^{-1/2}nm^{-1}$) and Φ_W represents the mean of the work functions of the tip and the sample ($\approx$ 3–5 eV).

Figure 3.36 schematically presents the principal components of an STM appartus. A sharp metal tip sweeps the sample surface at close distance. A feedback mechanism

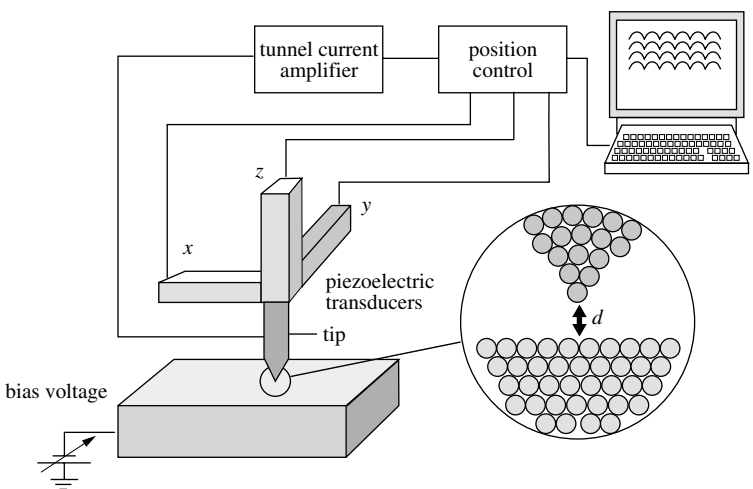


Figure 3.36 Schematic of scanning tunneling microscope (STM).

© 2007, first edition, EPFL Press

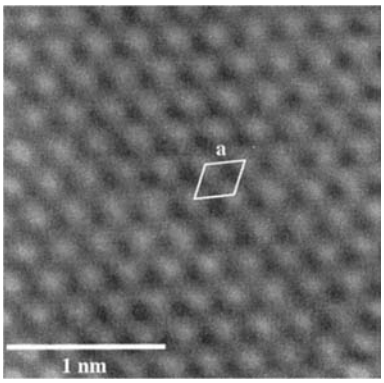


Figure 3.37 High resolution STM image of Ni [111] surface (ref. [10]).

keeps the tunneling current constant by adjusting the vertical distance between tip and sample surface. Sensors measure the vertical displacement of the probe as a function of its position along the *x* and *y* axes. The data are stored in a computer and serve to establish a topographic map of the surface using appropriate software. The resolution of an STM may approach 0.01 nm along the vertical axis and 0.2 nm along the horizontal axes.

In reality, the STM current depends on the local electron density around the probe. In order to obtain topographic information, all the surface atoms therefore should have identical electronic properties; in other words, they ought to be chemically indistinguishable. Under these conditions, the scanning tunneling microscope is capable of giving an accurate image of a surface at the atomic scale. In contrast to LEED or FIM, the STM does not require a vacuum; it can be operated in air or even in solution. On the other hand, because a tunnel current is measured, the use of the STM is restricted to the study of conducting surfaces. Figure 3.37 shows an STM

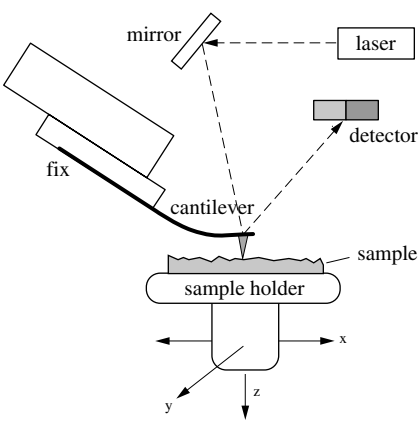


Figure 3.38 Working principle of an atomic force microscope (AFM).

© 2007, first edition, EPFL Press

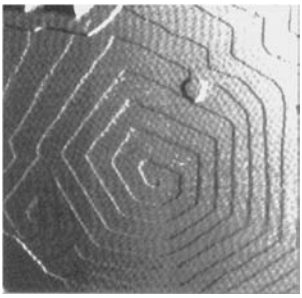


Figure 3.39 AFM image of copper surface produced by electrodeposition (ref. [12]).

image of the [111] surface of nickel. The distance between atoms agrees well that determined by X-ray diffraction [10].

The ***atomic force microscope***, or ***AFM***, developed in 1986 [11], is a local probe method similar as the STM. Instead of an electric tunneling current, one measures attractive or repulsive forces between the probe and the surface atoms. This allows the AFM to be used on insulating substrates. The probe, formed as a fine tip that can be either insulating or conducting is attached to a microscopic cantilever of small mass. In the so-called contact mode, as the tip approaches the surface the repulsive force leads to a deflection of the cantilever that can be monitored with a suitable detector. Figure 3.38 schematically shows an AFM using a laser detector. A feed back mechanism maintains a constant distance between the probe tip and the analyzed surface as a function of the measured force. Modern atomic force microscopes and scanning tunneling microscopes possess powerful image processing capabilites, that permit to present the results of probe scanning as a three-dimensional map.

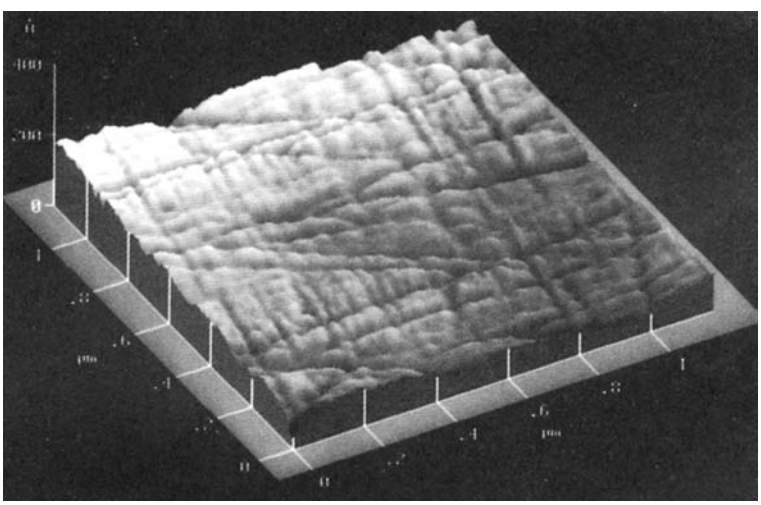


Figure 3.40 AFM image of mechanically polished stainless steel surface, showing the rugosity on the microscopic scale (adapted from ref. [8]).

© 2007, first edition, EPFL Press

Furthermore, both methods use an absolute height scale thus providing a quantitative measurement of roughness down to the atomic scale. Figure 3.39 shows the AFM image of a copper surface observed in electrodeposition [12]. The monatomic steps and the ledges form a growth spiral indicating that growth proceeded from a dislocation crossing the surface. When coupled with an appropriate detection system, the AFM is also capable of imaging rugosity at much larger scales. the AFM image of Figure 3.40 shows a mechanically polished stainless steel surface with characteristic polishing mark.

Characterization of surface topography at the microscopic scale

Surface topography at the microscopic scale is particularly important for engineers who deal with technical problems. Optical microscopy, scanning electron microscopy and stylus methods are most often used for characterizing surface structure and topography at the microscopic scale.

The ***scanning electron microscope*** or ***SEM*** is widely used for the observation metal surfaces. A tightly focused electron bean sweeps the sample surface, while a detector monitors the secondary electron flux emitted by the irradiated sample. The local electron flux depends on the angle created between the incident beam and the surface. Thus, an image is formed that reflects the surface topography. SEMs equipped with a field-emission electron source attain a spatial resolution down to some nanometer. On the other hand, the SEM also offers the option of operating at small magnification, yielding topographic information on a much larger scale. The SEM therefore is often used to study fracture surfaces (see Chapter 11). Indeed, compared to optical microscopy the SEM offers a much larger depth of field. Many SEMs are equipped with X-ray dispersive analysis that permits chemical analysis of samples with high spatial resolution.

Transmission electron microscopy or ***TEM*** can also be applied to study surface topography, but the procedure is tedious because it requires preparation of a carbon replica of the surface. More important is the use of TEM for the study of subsurface regions, by preparing a thin sample perpendicular to the surface and observing it in transmission. The recently developed *focussed ion beam technique (FIB)*, greatly facilitates sample preparation and is expected to lead to a wider use of the TEM observations of near surface regions. The FIB method uses a finely focussed high intensity ion beam for precisely cutting the sample surface and to produce a thin film that can be observed in transmission.

Optical microscopy, the most ancient method for studying surface topography, has two main limitations, a spatial resolution limited to about a micrometer, a limit set by the wavelength of visible light, and a restricted depth of field. Indeed, the higher the magnification one uses, the more accurately the observed feature must be located in the focus plane of the microscope lens. This limits the capability of optical microscopy to image rough surfaces. On the other hand, this apparent limitation can also be an advantage; it allows one to measure differences in height on a surface by focussing on different locations. An example would be the measurement of the depth

© 2007, first edition, EPFL Press

of corrosion pits. ***Optical interference microscopy*** is based on the superposition of two beams, one sent by a mirror serving as a reference, and the other by the sample surface. Interference fringes appear, separated by half wavelengths. Depending on its depth, rugosity such as a scratch will show up as a shifting of the interference fringes. The method gives a high vertical resolution (on the order of 0.01 μ), but the lateral resolution is limited. Furthermore, on rough surfaces the interpretation of interference fringes is difficult. Optical interference microscopy is sometimes used for checking the quality of polished surfaces.

Stylus methods are commonly used for measuring the rugosity of engineering surfaces. Typically, a diamond tip moves across a surface while a system records its vertical displacement. Instead of a diamond tip some instruments use a finely focussed laser beam, to avoid any possibility of scratching the surface under observation. A feedback mechanism adjusts the height of the laser tip based on the measurement of the intensity of the reflected light; when the surface is at the focal point of the lens, the intensity of the reflected light is at its maximum. Single scans provide a two-dimensional rugosity profile from which characteristic roughness parameters can be determined as described below. Using repeated scanning in the *x-y* plane, a three-dimensional surface map can be established using appropriate software. Stylus methods for measuring surface roughness thus yield similar information as the AFM, but on a larger measuring scale. They are therefore particularly well adapted to the characterization of rough surfaces.

To characterize the ***rugosity*** of technical surfaces different roughness parameters are used in practice [13]. The most commonly used parameter is the *average rugosity*, R_a, which represents the arithmetic average distance, in absolute value, between all profile positions and the average level of the surface.

$$R_a = \frac{1}{L} \int_0^L y_{(x)} \, dx \tag{3.74}$$

In (3.74), L is the measuring length, i.e., the distance covered by the probe, and $y(x)$ is the profile amplitude. In a similar fashion, the *root mean square rugosity* R_q characterizes the geometric average (root mean square) of the amplitude.

$$R_q = \left(\frac{1}{L} \int_0^L y_{(x)}^2 \, dx \right)^{1/2} \tag{3.75}$$

R_a and R_q are parameters related to the amplitude and as such do not give any indication about the wavelength of the rugosity nor about the shape of the surface features. Other parameters, called lateral separation parameters, take into account these aspects, but they are less commonly used.

The absolute value of R_a and R_q depends on the interval length considered for the measurement. This is illustrated by Figure 3.41 for titanium surfaces polished either mechanically or electrochemically [14]. At small interval length the R_q for the electropolished surface is smaller than for the mechanically polished surface, reflecting the absence of polishing scratches. At an interval length exceeding 300 micrometer

© 2007, first edition, EPFL Press

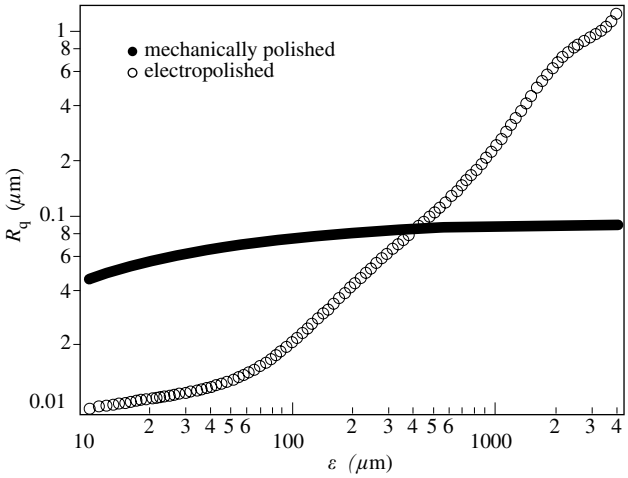


Figure 3.41 Effect of measurement interval length in the value of the mean roughness R_q of electrochemically or mechanically polished titanium surfaces (after ref. [14]).

the opposite behavior is found, because a slightly non-uniform current distribution during electropolishing led to a small curvature of the originally flat sample surface that is reflected in R_q.

3.5 ELECTRODE-ELECTROLYTE INTERFACE

3.5.1 Electric double layer

A metal crystal can be thought of as the combination of positive ions occupying the positions of a lattice and mobile electrons that behave like a gas. On the whole, the crystal is electrically neutral, i.e., the number of positive and negative charges is identical. Despite their kinetic energy, the electrons are not capable to escape more than 1–2 nm from the crystal surface because of the attractive force exerted by the positive charges. Figure 3.42 represents the electron distribution at a metal surface exposed to vacuum. A thin zone, in immediate proximity to the solid, carries a

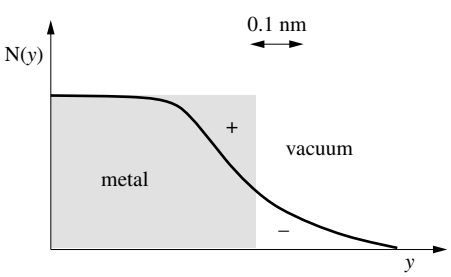


Figure 3.42 Variation of the electron density with distance at a metal surface.

© 2007, first edition, EPFL Press

negative charge, while a second thin layer located below the surface has a positive charge.

Quantum mechanics offers a more realistic description of this behavior. At the surface of the metal, a potential barrier retains the electrons. However, there is a certain probability that electrons tunnel through the barrier and thus can exist in the immediate vicinity of the surface. The electron distribution outside the solid depends on, among other parameters, the electronic properties of the solid and the characteristics of the potential barrier.

Charge distribution at the metal-electrolyte interface

In a similar way as described above for the metal-vacuum interface one observes a charge separation when a metal surface is contact with an electrolyte. The charge distribution at the metal-electrolyte interface depends on numerous factors: the electronic properties of the solid, the adsorption of water molecules and of hydrated cations, and the chemisorption of anions. The zone of charge separation near the interface is called the ***electric double layer***, or simply the ***double layer***. Depending on the applied potential, the charge of the metal can either be positive or negative relative to the electrolyte. The structure of the double layer therefore varies with the potential and with the chemical nature of the ions present. However, in all cases the sum of the positive charges in the whole of the double layer equals that of the negative charges. Figure 3.43 gives a schematic representation of the double layer region, for the specific case of excess negative charge at the metal surface.

Water is a dipolar molecule. Therefore, when it adsorbs, it orients itself according to the charges located at the surface. Dissolved cations are generally hydrated. Because of this, they are unable to approach the surface any closer than the distance defined by the radius of their hydration shell. Anions are not hydrated in general. Certain anions such as chloride are capable of chemisorbing onto the surface, replacing adsorbed

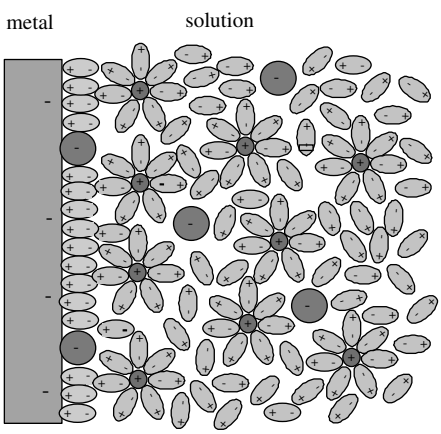


Figure 3.43 Electrical double layer at metal-electrolyte interface in the presence of chemisorbed anions.

© 2007, first edition, EPFL Press

water molecules. Interestingly, this may take place even if the metal is negatively charged.

Solid metal surfaces contain numerous defects such as atomic steps, point defects and dislocations and grain boundaries that can influence double layer properties. Most basic experiments on the structure of the electric double layer are therefore carried out on liquid electrodes, mostly mercury. This also avoids problems related to rugosity and crystal orientation. Experiments with polycrystalline solids yield an average value for the double layer properties of the electrode-solution interface.

Electric analog to the double layer

By definition, the double layer is that region of the interface where a separation of charges is observed. Electric neutrality requires that the absolute value of the charge on the metal per unit surface, q_m, is equal to that of the adjacent electrolyte, q_s, and that the signs of the charge be opposite. This gives:

$$q_m + q_s = 0 \tag{3.76}$$

The electrical behavior of a metal-electrolyte interface can thus be compared to a capacitor connected in parallel to a resistance (Figure 3.44). The latter is called the ***charge transfer resistance***, or more simply the *transfer resistance*. In reality, the charge transfer resistance varies with the current and therefore does not behave like an ohmic resistance (cf. Chap. 4). For small potential or current variations the analogy remains valid, however.

The electric circuit shown in Figure 3.44, called an ***equivalent circuit***, shows that when the double layer is charged a current can cross the interface. This current is called charge transfer current, or ***faradaic current***, because according to Faraday's law the current flow is associated with an electrochemical reaction at the electrode-electrolyte interface. In the absence of reactive species and with metals exhibiting a large hydrogen overvoltage (see Chap. 4) the electrode potential can be varied over a wide range without producing a faradaic current. Electrode-electrolyte systems that possess this property are called ***ideally polarizable electrodes***. A classical case are

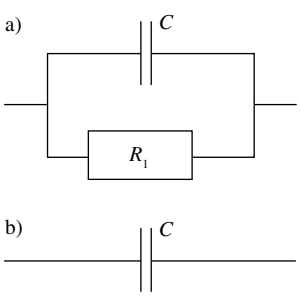


Figure 3.44 Equivalent circuit of metal-electrolyte interface for (a) an electrode with charge transfer reaction, showing the double layer capacitance and the charge transfer resistance and for (b) an ideally polarizable electrode with infinite charge transfer resistance.

© 2007, first edition, EPFL Press

mercury electrodes immersed into solutions of inert salts, such as NaCl and NaF. In addition to being a liquid at room temperature, mercury is therefore particularly well suited for the study of double layer structure. The transfer resistance of an ideally polarizable electrode being infinite, the equivalent circuit contains only a capacitor.

Helmholtz model.

Different theoretical models have been developed to describe the electrical properties of the double layer. The simplest of these, proposed by Helmholtz (Figure 3.45), assumes that the positive and negative charges are located in two planes separated by a distance L_H of about 0.2 to 0.3 nm, which corresponds to the minimum distance separating hydrated ions from the electrode surface.

According to the Helmholtz model, the double layer is analogous to a planar capacitor, whose capacity can be calculated by the equation:

$$C_H = \frac{dq}{d\Delta\Phi} = \frac{\varepsilon \cdot \varepsilon_0}{L_H} \tag{3.77}$$

where ε represents the dielectric constant of the adsorbed water, ε_0 is the permittivity constant, and $q = q_m$ is the charge density (C m^{-2}). The quantity $\Delta\Phi$ represents the potential difference between the metal and the electrolyte.

$$\Delta\Phi = \Phi_m - \Phi_{\sigma,b} \tag{3.78}$$

According to equation (3.77), the capacitance of the double layer depends neither on the potential nor on the concentration of the electrolyte. Taking typical values of $\Delta\Phi$ = 0.5 V, L_H = 0.5 nm and ε = 78, we find a capacity of C_H = 0.7 F m^{-2} = 70 μF cm^{-2}. This value agrees reasonably well with measurements taken with mercury electrodes. And yet, the real dielectric constant of water in the double layer certainly

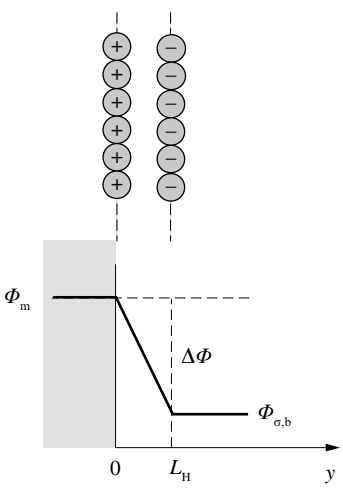


Figure 3.45 Helmholtz model of electric double layer.

© 2007, first edition, EPFL Press

differs from the indicated value, varying as a function of the orientation of the adsorbed water molecules. If this orientation is potential dependent, the dielectric constant of the adsorbed water and the capacitance should also vary with the potential. It is still possible to estimate the electric field in the double layer: $\Delta\Phi/L_H = 10^9\ \mathrm{Vm}^{-1}$. Clearly, the electric field is very strong inside the Helmholtz double layer.

Gouy-Chapman Model

In dilute electrolytes, the concentration of ions in the electrolyte has been found to influence the capacity of the double layer. This behavior contradicts the Helmholz model, which does not contain concentration. The origin of the concentration dependence of the double layer capacity lies in the thermal motion of the ions. In fact, the ions do not align themselves into a stable position, as assumed in the Helmholtz model, but are spread out according to a Boltzmann distribution within a zone close to the surface, called the ***diffuse double layer*** or the ***Gouy-Chapman layer*** (Figure 3.46). By defining a charge density in the diffuse layer ρ_z, a bulk concentration $c_{i,b}$ of ions i in the electrolyte and an ion charge z_i, the Boltzmann distribution can be written as:

$$\rho_z = F \sum z_i\, c_{i,b} \exp\left(-\frac{z_i F \Phi}{RT}\right) \tag{3.79}$$

where $\Phi = f(y)$ represents the potential within the double layer with respect to that in the bulk electrolyte.

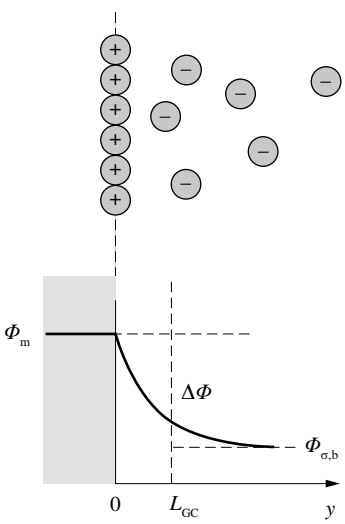


Figure 3.46 Gouy-Chapman model of electric double layer.

© 2007, first edition, EPFL Press

$$\Phi = \Phi_\sigma - \Phi_{\sigma,\mathrm{b}} \tag{3.80}$$

To simplify, we set the potential in the electrolyte $\Phi_{\sigma,\mathrm{b}} = 0$ and the subscript σ is omitted.

For a ***binary symmetric electrolyte***, we have by definition (the subscripts + and – stand for the cation and anion, respectively) $z_+ = -z_- = z$ and $c_+ = c_- = c$. Equation (3.79) then becomes:

$$\rho_z = z\,F\,c_\mathrm{b}\left[\exp\left(\frac{zF\Phi}{RT}\right) - \exp\left(-\frac{zF\Phi}{RT}\right)\right] \tag{3.81}$$

The Poisson equation provides a relationship between the charge density ρ_z and the potential Φ:

$$\varepsilon \cdot \varepsilon_0 \cdot \frac{\mathrm{d}^2\Phi}{\mathrm{d}y^2} = -\rho_z \tag{3.82}$$

With (3.81), and by setting

$$\frac{\mathrm{d}^2\Phi}{\mathrm{d}y^2} = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\Phi}\left(\frac{\mathrm{d}\Phi}{\mathrm{d}y}\right)^2 \tag{3.83}$$

the integration of (3.82) gives the potential gradient [15]:

$$\frac{\mathrm{d}\Phi}{\mathrm{d}y} = \left(\frac{8\,RT\,c_\mathrm{b}}{\varepsilon \cdot \varepsilon_0}\right)^{1/2} \sinh\left(\frac{z\,F\,\Phi}{2\,RT}\right) \tag{3.84}$$

According to Gauss's law, the electric charge Q bounded by the surface A is equal to the electric flux that passes through it. This flux is proportional to the integral over the surface of the potential gradient.

$$Q = \varepsilon \cdot \varepsilon_0 \oint \nabla\,\Phi\,\mathrm{d}A \tag{3.85}$$

For a planar electrode, Gauss' law is written as:

$$q_\mathrm{s} = \varepsilon \cdot \varepsilon_0 \left.\frac{\mathrm{d}\Phi}{\mathrm{d}y}\right|_{y=0} \tag{3.86}$$

where $q_\mathrm{s} = -q_\mathrm{m}$ represents the charge of the electrode per unit surface area. Equation (3.83), with the additional condition

$$y = 0; \quad \Phi = \Phi_\mathrm{m} \tag{3.87}$$

allows us to calculate the charge of the double layer, $q = q_\mathrm{m}$:

© 2007, first edition, EPFL Press

$$q = \varepsilon \cdot \varepsilon_0 \left(\frac{8\,RT\,c_b}{\varepsilon \cdot \varepsilon_0} \right)^{1/2} \sinh\left(\frac{z\,F\Delta\Phi}{2\,RT} \right) \tag{3.88}$$

In this equation, we have set $\Delta\Phi = \Phi_m - \Phi_{\sigma,b} = \Phi_m$. Differentiation with respect to the potential gives us the capacity of the double layer as a function of $\Delta\Phi$:

$$C_{GC} = \frac{dq}{d\Delta\Phi} = \varepsilon \cdot \varepsilon_0 \left(\frac{2\,z^2\,F^2\,c_b}{\varepsilon\,\varepsilon_0 RT} \right)^{1/2} \cosh\left(\frac{z\,F\Delta\Phi}{2\,RT} \right) \tag{3.89}$$

Equation 3.90 defines the thickness of the diffuse double layer L_{GC}

$$L_{GC} = \left(\frac{\varepsilon\,\varepsilon_0\,RT}{2\,z^2\,F^2\,c_b} \right)^{1/2} \tag{3.90}$$

The capacity of the diffuse double layer therefore has the value:

$$C_{GC} = \frac{\varepsilon \cdot \varepsilon_0}{L_{GC}} \cosh\left(\frac{z\,F\Delta\Phi}{2\,RT} \right) \tag{3.91}$$

According to (3.90), the thickness of the diffuse double layer decreases as the concentration of the electrolyte increases. Taking a value of the dielectric constant of $\varepsilon = 78$, an ion charge of $z = 1$, a temperature of 298 K and a concentration of 10^{-4} mol l^{-1}, the thickness of the diffuse double layer is determined to be $L_{GC} = 30$ nm, which is significantly larger than the value proposed by the Helmholtz model.

The integration of (3.83) allows us to calculate the variation of the potential as a function of the distance from the surface. To simplify the integration, we set: $\sinh(x) \approx x$. The result (3.91) shows that the potential in the diffuse layer varies exponentially.

$$\Phi = \Phi_m \exp\left(-\frac{y}{L_{GC}} \right) \tag{3.92}$$

Gouy-Chapman theory can also be applied to insulating materials and to colloids. In this case, the potential corresponding to y = 0 is called the ***zeta potential***. The zeta potential plays a crucial role for the stability of colloids and for electro-osmotic phenomena in porous insulating materials.

Stern Model

The Stern Model is a combination of the Helmholtz and Gouy-Chapman models (Figure 3.47). The potential difference between the metal and the solution is comprised of two terms: $\Delta\Phi_H$, due to the compact Helmholtz layer and $\Delta\Phi_{GC}$, due to the diffuse Gouy-Chapman layer.

© 2007, first edition, EPFL Press

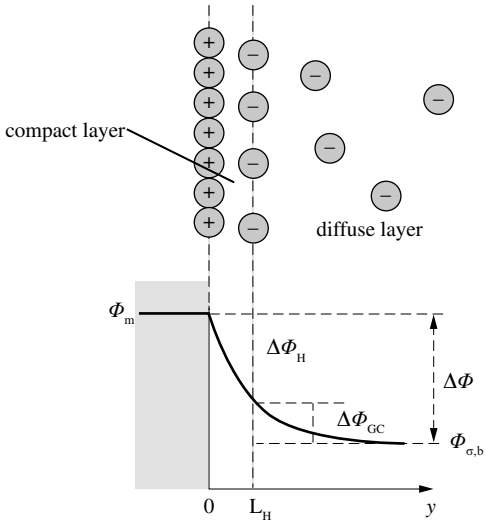


Figure 3.47 The Stern model.

$$\Delta\Phi = \Delta\Phi_H - \Delta\Phi_{GC} \tag{3.93}$$

By analogy to a circuit consisting of two capacitors in series, the following equation gives the total capacity of the double layer:

$$\frac{1}{C} = \frac{1}{C_H} + \frac{1}{C_{GC}} \tag{3.94}$$

It follows from this expression that the smaller of the two capacities essentially determines the overall value of the double layer capacity.

Figure 3.48 shows experimental values for the double layer capacity as a function of the applied potential for a mercury electrode in NaF solutions at different concentrations [16]. The curves for low concentrations (0.001 M and 0.01 M) show a pronounced minimum, which is in agreement with the hyperbolic cosine dependence in equation (3.90). At these concentrations the effect of the diffuse layer dominates. As the concentration increases, the minimum disappears because the Helmholtz layer governs the overall capacity.

The Stern model provides a reasonable description of the electric behavior of the metal-electrode interface for electrochemical systems, but it does not allow us to explain all experimental observations. In particular, it does not offer a satisfactory explanation for the influence of crystal orientation on the capacity of the double layer, nor for the effect of the chemical nature of anions. Figure 3.49 [17] shows the capacity of the double layer of a monocrystalline silver electrode, with three different orientations. The resulting curves are similar, but they are displaced along the axis of the potential.

© 2007, first edition, EPFL Press

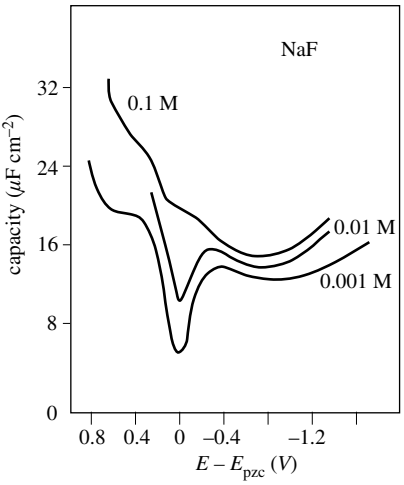


Figure 3.48 Double layer capacity of a mercury electrode in NaF solutions of different concentrations, plotted as a function of potential with respect to potential of zero charge. Electrolyte is NaF at 0.1, 0.01, and 0.001 M (adapted from ref. [16]).

Newer models of the electrode-electrolyte interface, based on quantum mechanics, will not be presented here. By taking account of the electron distribution near the surface, these models can give, in principle, an explanation for the influence of crystal orientation, but at this time they do not offer decisive insight into the variety of chemical effects observed.

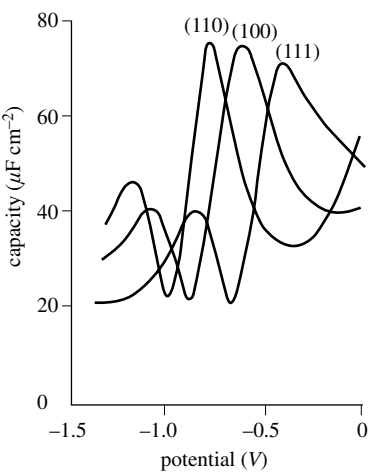


Figure 3.49 Double layer capacity of silver monocrystal of different orientation plotted as a function of potential measured against the saturated calomel reference electrode. Electrolyte is 0.01 M NaF (adapted from ref. [17]).

© 2007, first edition, EPFL Press

3.5.2 Experimental determination of the double layer charge

Electrocapillary curve and potential of zero charge

In chemical systems, the Gibbs equation (3.23) describes the variation of the surface tension as a function of the chemical potential of the adsorbed species. In electrochemical systems, the surface tension depends also on the prevailing electrode potential and can be calculated by the ***Lippmann equation***. For a binary electrolyte, which, by definition, has only a single type of cation and anion, the Lippmann equation has the following form [18]:

$$-\mathrm{d}\gamma = \frac{\Gamma_\mathrm{i}}{\nu_\mathrm{i}}\mathrm{d}\mu + q\,\mathrm{d}E \tag{3.95}$$

where Γ_i is the excess of cations or anions i inside the double layer; μ is the chemical potential of the dissolved salt; ν_i is the stoichiometric coefficient for the ions i (ν_i = 1, for a symmetric salt); E is the electrode potential, and q the charge of the double layer ($q = q_\mathrm{m} = -q_\mathrm{s}$).

A graphical representation of the surface tension as a function of the electrode potential is called ***electrocapillary curve***. Figure 3.50 [19] shows several examples of such curves, measured with a mercury electrode immersed in NaF solutions of different concentration. The slope gives the charge of the double layer.

$$q = -\left(\frac{\partial\gamma}{\partial E}\right)_\mu \tag{3.96}$$

The charge density q results from an excess of anions or cations within the double layer:

$$q = q_\mathrm{m} = -q_\mathrm{s} = -z_\mathrm{i}\cdot F\cdot\Gamma_\mathrm{i} \tag{3.97}$$

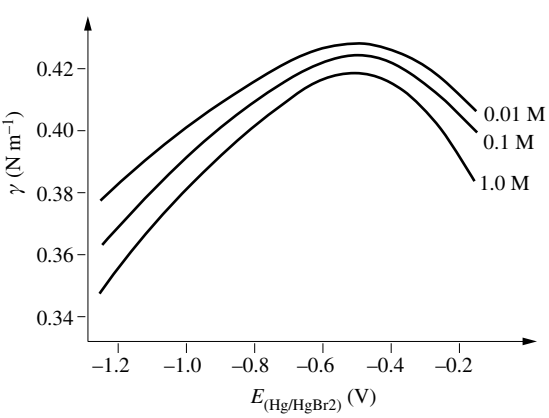


Figure 3.50 Variation of surface tension of mercury electrode as a function of applied potential in solutions of NaBr of different concentration (adapted from ref. [19]).

© 2007, first edition, EPFL Press

In (3.97), the charge z_i of cations is positive, and that of anions negative. In Figure 3.50, potentials greater than that which corresponds to the maximum of the curve indicate an excess of anions in the double layer; at potentials lower than the maximum, cations are in excess. At the maximum of the electrocapillary curve, the charge of the electrode is zero, there is no ionic excess. The potential that corresponds to this situation is called the ***potential of zero charge***.

Capacity measurements

The surface tension of solid electrodes is not easily measured. Capacity measurements are a more convenient way to determine the potential of zero charge. The differential capacity of an electrode is defined by:

$$C = \left(\frac{\partial q}{\partial E}\right)_{\mu} \tag{3.98}$$

At the potential of zero charge, the double layer capacity exhibits a minimum.

$$E = E_{\text{pzc}}; \quad \frac{\partial C}{\partial E} = 0 \tag{3.99}$$

In order to determine the electrode charge as a function of a given potential E', equation (3.97) is integrated between the potential of zero charge (E_{pzc}) and the potential E'.

$$q = \int_{E_{\text{pzc}}}^{E'} C \, \mathrm{d}E \tag{3.100}$$

Figure 3.51 schematically shows how the charge of the double layer can be determined by (a) electrocapillary measurements and (b) by capacitance measurements. We may keep in mind that an anionic excess in the double layer corresponds to a positive charge at the metal and therefore to a positive value of q.

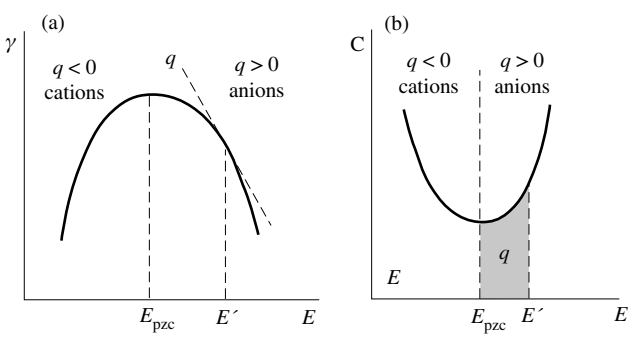


Figure 3.51 Experimental determination of double layer charge as a function of potential by (a) the meaurement of the surface tension and (b) the measurement of the double layer capacitance.

© 2007, first edition, EPFL Press

In the absence of chemical adsorption, the value of the potential of zero charge depends on the electronic properties of the metal and on its crystal orientation. Certain research groups have observed a linear relationship between the potential of zero charge and the work function. In practice, because of adsorption, the potential of zero charge of a metal is strongly influenced by its environment.

Specific adsorption

According to the classical models of Helmholtz, Gouy-Chapman and Stern, the charge resulting from the presence of excess ions in the double layer is balanced by an equal charge of opposite sign in the metal. Under certain conditions, it turns out that chemically adsorbed anions compensate part of the ionic charge (see Figure 3.43):

$$q_m + q_s + q_{ads} = 0 \tag{3.101}$$

Here q_{ads} is the charge of chemisorbed anions. The layer parallel to the interface, defined by the point of nearest approach of hydrated ions, is called the *outer* Helmholtz plane; and the one at corresponding to the position of chemisorbed anions is called the *inner* Helmholtz plane. Ions located in the inner layer are in direct contact with the surface, without the interposition of water molecules. This phenomenon, referred to as chemical adsorption or ***specific adsorption***, particularly concerns certain types of anions, the most important from a corrosion point of view being chloride. In order to study specific adsorption, the experimental charge of the double layer is compared to the predictions of the Stern model and the difference is attributed to specific adsoption.

The discussed models have been developed for liquid electrodes, as mentioned previously. On solid surfaces, due to the presence of atomic steps and crystal defects different atomic sites exist which differ in energy. Specific adsorption then is expected to occur preferentially at certain sites rather than uniformly. Site specific adsorption phenomena can in principle be studied with local probe methods such as the STM, but their discussion is beyond the scope of this book.

3.5.3 Semiconductor-electrolyte interface

Most engineering metals are covered by a thin oxide film that provides protection against corrosion. Many of these films are semiconductors or insulators. Therefore, the properties of the semiconductor-electrolyte interface are of considerable interest for metal corrosion.

Basic concepts.

Semiconductors differ from metals by their electronic band structure. In metals, the valence band is only partially occupied by electrons and overlaps with the conduction band (Figure 3.52). Valence electrons can easily access non-occupied states and move under the influence of an applied electric field. Semiconductors and

© 2007, first edition, EPFL Press

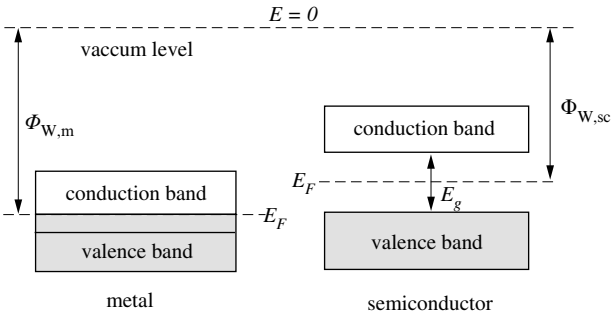


Figure 3.52 Band structure of a metal and of a semiconductor.

insulators have, in contrast, an entirely occupied valence band (at zero Kelvin) and an empty conduction band. A ***band gap*** separates the two energy bands. The band gap energy, E_g, is equal to the energy difference between the upper edge of the valence band and the lower edge of the conduction band.

At temperatures above absolute zero, some electrons will have enough thermal energy to move from the valance band to the conduction band. Each electronic transition leaves a "hole" in the valence band. These holes have an effective positive charge, are mobile, and contribute to the conduction of electrical charge, along with the electrons. Insulators differ from semiconductors in that their band gap is wider (4–5 eV compared to 2–3 eV for semiconducting oxides). In insulators the valence electrons are no longer able to access the conduction band and, as a result, electron movement is not induced by the application of an electric field.

The ***Fermi level*** characterizes the electron energy in a solid. In metals the Fermi energy equals the chemical potential of the electrons (see Chapter 2). It corresponds to the highest energy level occupied by electrons at zero Kelvin. At higher temperature, the distinction between occupied and non-occupied levels spreads across an energy interval that lies above and below the Fermi level. At the Fermi level, the occupation probability is just 1/2.

The Fermi level of semiconductors is located in the band gap. For ***intrinsic semiconductors***, the electron concentration c_n in the conduction band is equal to the hole concentration c_p in the valence band. The equilibrium constant can be written as:

$$c_n \cdot c_p = c_{pn}^2 = K \tag{3.102}$$

We call c_{pn} the *intrinsic carrier density*. The equilibrium constant can be expressed by the formula

$$K = K_0 \exp(-E_g/kT) \tag{3.103}$$

where K_0 is a constant.

The Fermi level in an intrinsic semiconductor is located close to the center of the band gap, and it is called the intrinsic energy level E_{int}.

© 2007, first edition, EPFL Press

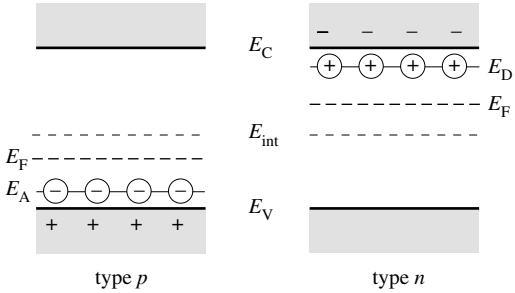


Figure 3.53 Fermi level in p-type and n-type semiconductors.

In ***extrinsic semiconductors***, unlike the precedent case, the numbers of positive and negative charge carriers differ. In ***n-type*** semiconductors, the electron concentration is much greater than that of the holes: $c_n >> c_p$. A ***p-type*** semiconductor corresponds to the opposite case: $c_n << c_p$. In silicon, these properties result from the doping with an electron acceptor (boron) or donor (phosphorus). A similar effect is observed in oxides with an excess or a deficiency of oxygen ions. For example, the oxides ZnO, TiO_2, V_2O_5, MO_3 are n-type and the oxides Cr_2O_3, NiO, Cu_2O are p-type semiconductors (see Chapter 9).

Figure 3.53 shows the influence of doping. An acceptor introduces new quantum states into the forbidden gap, close to the valence band. These additional energy levels are easily accessible to the electrons. An electron transfer into a new level leaves behind a hole in the valence band. Thus, the concentration of holes c_p is equal to the acceptor concentration, c_A.

$$c_p = c_A \tag{3.104}$$

In an analogous manner, the presence of donors introduces new energy levels near the conduction band. Electrons from the donor that move into the conduction band become negative charge carriers. Their concentration c_n is equal to the donor concentration c_D.

$$c_n = c_D \tag{3.105}$$

The presence of dopants alters the Fermi level, which is located closer to the conduction band in n-type semiconductors and closer to the valence band in p-type semiconductors.

Space-charge layer

The double layer at a semiconductor-electrolyte interface differs from that found at metal-electrolyte interfaces in that the charges in the solid are distributed over a certain thickness, the ***space charge layer***.

The space-charge layer in semiconductors is comparable to the diffuse double layer in electrolytes that forms at a low concentration. Indeed, the concentration of

© 2007, first edition, EPFL Press

charge carriers in a semiconductor and in a dilute electrolyte solution are often similar. For example, a doping of 10^{17} charge carriers per cm^3 in silicon corresponds to a concentration of 1.7×10^{-4} mole charge carrier/dm^3, comparable to that of a dilute electrolyte solution.

Mott-Schottky equation

The formation of a space-charge layer leads to a positive or negative bending of the energy bands near the surface (Figure 3.54). As a result, a potential difference is established in the semiconductor, between the bulk $\Phi_{sc,b}$ and the surface $\Phi_{sc,s}$:

$$\Delta\Phi_{sc} = \Phi_{sc,b} - \Phi_{sc,s} \tag{3.106}$$

By analogy with the charge of the Gouy-Chapman layer, we can calculate q_{sc} for the space-charge layer. For intrinsic semiconductors this yields:

$$q_{sc} = \frac{2\varepsilon\varepsilon_0 RT}{FL_{sc}} \sinh\left(\frac{F\Delta\Phi_{sc}}{2RT}\right) \tag{3.107}$$

where L_{sc}, called the *Debye length*, represents the thickness of the space charge layer and is given by:

$$L_{sc} = \left(\frac{\varepsilon\varepsilon_0 RT}{2F^2 c_{pn}}\right)^{1/2} \tag{3.108}$$

Here, c_{pn} is the carrier bulk concentration of the semiconductor.

The capacitance of the space charge layer is equal to:

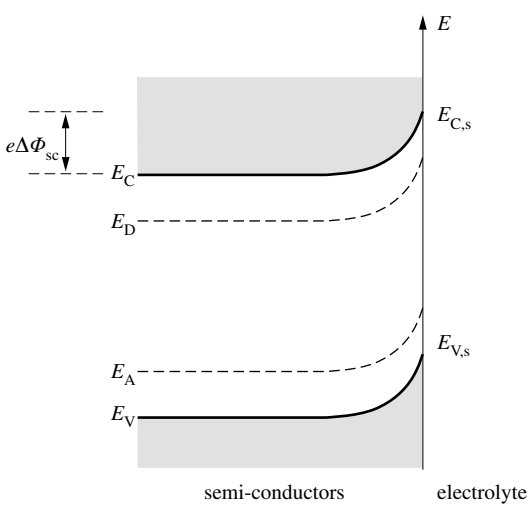


Figure 3.54 Bending of energy bands at the surface of a semiconductor electrode.

© 2007, first edition, EPFL Press

$$C_{sc} = \frac{dq_{sc}}{d\Delta\Phi_{sc}} = \frac{\varepsilon\varepsilon_0}{L_{sc}} \cosh\left(\frac{F\Delta\Phi_{sc}}{2RT}\right) \tag{3.109}$$

The presence of dopants complicates the greatly calculation of the space charge. The result has the form [20]

$$q_{sc} = \frac{2\varepsilon\varepsilon_0 RT}{FL_{sc}} \psi(\lambda, \Delta\Phi_{sc}) \tag{3.110}$$

where $\psi(\lambda, \Delta\Phi_{sc})$ is a function of $\Delta\Phi_{sc}$ and of the ratio $\lambda = c_{pn}/c_n = c_p/c_{pn}$, which characterizes the extent of doping. A space-charge layer in which the concentration of the majority carriers is lower than in the bulk is called ***depletion layer***. For these conditions the Mott-Schottky equation (3.111) describes the variation of the capacitance as a function of potential.

$$\frac{1}{C_{sc}^2} = \left(\frac{2L_{sc}}{\varepsilon\varepsilon_0}\right)^2 \left(\frac{F\Delta\Phi_{sc}}{RT} - 1\right) \tag{3.111}$$

$$L_{sc} = \left(\frac{\varepsilon\varepsilon_0 RT}{2F^2 c_n}\right)^{1/2} \tag{3.112}$$

For a *p*-type semiconductor, c_n is replaced by c_p in (3.112).

Influence of the electrolyte

The potential difference at the semiconductor-electrolyte interface is the sum of three terms: $\Delta\Phi_{sc}$ rising from the space-charge layer in the semiconductor; $\Delta\Phi_H$ from the Helmholtz layer; and $\Delta\Phi_{GC}$ from the Gouy-Chapman layer (Figure 3.55).

$$\Delta\Phi = \Delta\Phi_{sc} + \Delta\Phi_H + \Delta\Phi_{GC} \tag{3.113}$$

The electrode potential E is given by $\Delta\Phi$ and an additional constant whose value depends on the reference electrode.

$$E = \Delta\Phi + \text{constant} \tag{3.114}$$

The capacity of the double layer of a semiconducting electrode can be associated with three capacitors connected in series:

$$\frac{1}{C} = \frac{1}{C_{sc}} + \frac{1}{C_H} + \frac{1}{C_{GC}} \tag{3.115}$$

Most often the $1/C_{sc}$ term dominates and one simply gets:

$$\frac{1}{C} \approx \frac{1}{C_{sc}} \tag{3.116}$$

© 2007, first edition, EPFL Press

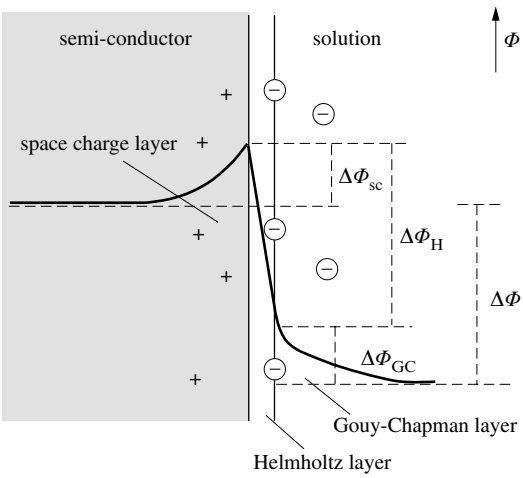


Figure 3.55 Double layer at semiconductor-electrolyte interface.

Easily measurable, the capacity of a semiconducting electrode provides direct information about the space-charge layer. Figure 3.56 shows the experimental values for the capacitance of a ZnO electrode, an *n*-type semiconductor [21]. The data are presented in the form of a Mott-Schottky diagram. From the slope of the straight line, the concentration of the majority carriers can be determined. In this case, we have

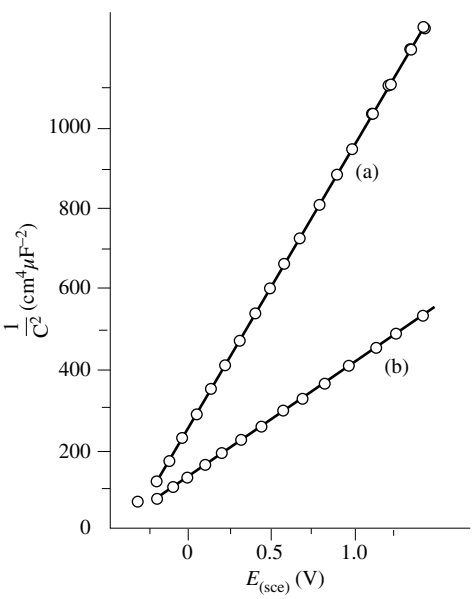


Figure 3.56 Mott-Schottky plot of ZnO electrodes in 1M KCl. Oxide conductivites are (a) 0.59 Ω^{-1} cm^{-1} and (b) 1.79 Ω^{-1} cm^{-1}.

© 2007, first edition, EPFL Press

$n_o = 6 \times 10^{24}$ m^{-3}. The intersection of the straight line with the abscissa gives the ***flat band potential***, E_{FB}, measured with respect to a calomel electrode. The flat band potential corresponds to the electrode potential for which no bending of the energy bands is induced at the semiconductor surface: $\Delta\Phi_{sc} = 0$.

Surface states

The models presented up to this point only take electrostatic effects into account. And yet, on semiconductor-electrolyte interfaces, chemical interactions can take place between adsorbed species and the solid phase. These interactions create new, distinct energy levels, located at the surface, that influence the space-charge layer and the potential difference $\Delta\Phi_{sc}$. Quantification of these phenomena is difficult, however, and beyond the scope of this discussion.

3.5.4 Chemical equilibria at the oxide-electrolyte interface

Thin semiconducting oxide films play a particularly important role in corrosion. In this section we therefore briefly discuss chemical interactions between an oxide and an electrolyte, in particluar acid-base reactions.

An oxide surface in contact with an aqueous solution will hydrate; in other words it will acquire -OH groups. These groups can act as either an acid or a base, by donating or accepting a proton. Depending on the pH of the electrolyte, the oxide surface will thereby acquire a negative or positive charge (Figure 3.57). The pH value that corresponds to zero surface charge is called the ***isoelectric point*** or the ***point of zero charge***. Figure 3.58 shows numerical values of the isoelectric point of different oxides, taken from reference [22]. These values give an indication of the acid-base character of different oxides, but they are not precise. Indeed values reported in the literature vary considerably, because they are affected, among other parameters, by the nature and concentration of the electrolyte and by the presence of impurities [23].

The acid-base reactions of oxide surface groups can be described with the same formalism used for reactions in homogeneous phase. This gives for the reaction of a surface hydroxyl group with a proton, taking place at pH below the isoelectric point:

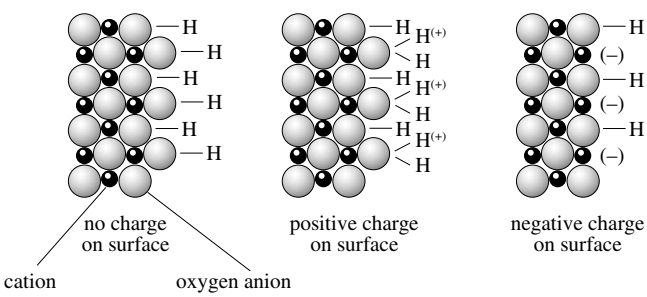


Figure 3.57 Oxide-electrode interface: (a) uncharged oxide; (b) positively charged oxide; (c) negatively charged oxide.

© 2007, first edition, EPFL Press

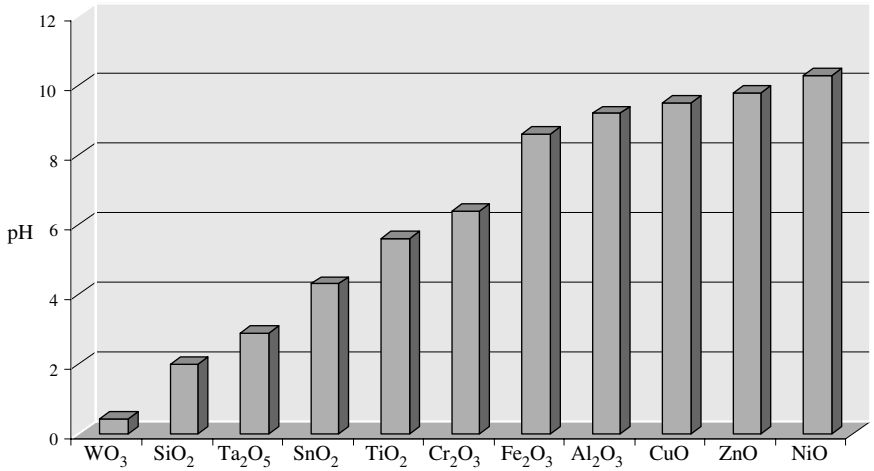


Figure 3.58 Numerical values of the isoelectric points of different oxides (ref. [22]).

$$\text{S–OH} + \text{H}^+ = \text{S–OH}_2^+ \tag{3.117}$$

The equilibrium constant of the reaction is written as:

$$K_1 = \frac{\left\{\text{S–OH}_2^+\right\}}{\{\text{S–OH}\} \cdot a_{\text{H}^+}} \tag{3.118}$$

where $a_{\text{H}+}$ represents the proton activity, and $\{\text{S–OH}_2^+\}$ and $\{\text{S–OH}\}$ are the activities of the surface groups $–\text{OH}_2^+$ and –OH.

Likewise, at pH above the isoelectric point, we find for a reaction of a hydroxyl group with an OH^- ion (deprotonation reaction):

$$\text{S–OH} + \text{OH}^- = \text{S–O}^- + \text{H}_2\text{O} \tag{3.119}$$

$$K_\text{b} = \frac{\left\{\text{S–O}^-\right\}}{\{\text{S–OH}\} \cdot a_{\text{OH}^-}} \tag{3.120}$$

With $K_{\text{H2O}} = a_{\text{H}+}\, a_{\text{OH}-}$ and $K_2 = K_\text{b}\, K_{\text{H2O}}$ we obtain:

$$K_2 = \frac{\left\{\text{S–O}^-\right\} \cdot a_{\text{H}^+}}{\{\text{S–OH}\}} \tag{3.121}$$

For example, for the hydrated oxide α-FeOOH in contact with a solution of 0.1 M NaClO_4, values of log K_1 = 6.4 and log K_2 = –9.25 have been reported [24] for low concentrations of the ionic surface groups.

© 2007, first edition, EPFL Press

Certain electrolyte ions can adsorb onto the oxide surface by taking the place of a proton or hydroxyl group; this phenomenon is particularly important for anions. The adsorption equilibrium is written as:

$$\text{S–OH} + A^- = \text{S–}A + \text{OH}^- \tag{3.122}$$

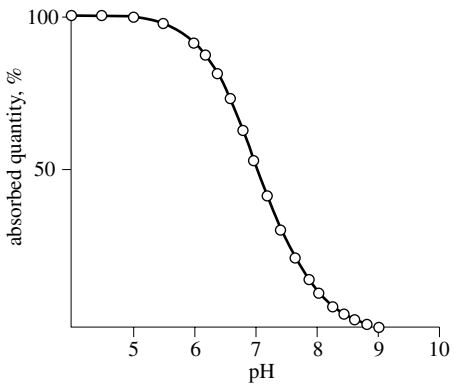


Figure 3.59 Effect of pH on adsorption of F^- ions on a γ-Al_2O_3 surface (adapted from ref. [24]).

The adsorption of an anion frees an OH^- ion. The reaction should therefore be favored by a low pH. This reasoning is in agreement with the results presented in Figure 3.59, which shows the variation of the fluoride ion coverage of an alumina surface (Al_2O_3) as a function of pH. Similarly, the adsorption of cations should be favored by a high pH because protons are liberated in the process.

General bibliography

G. A. Somorjai, *Principles of Surface Chemistry*, Prentice Hall, Englewood Cliffs, N.J. (l972) 283 pp.

A. W. Adams, A. P. Gast, *Physical Chemistry of Surfaces*, 6th edition, Wiley Interscience, New York (1997) 784 pp.

J. C. Vickerman, ed., *Surface Analysis, The Principle Techniques*, John Wiley and Sons Ltd, Chichester UK (1997) 457 pp.

P. Marcus, F. Mansfield, *Analytical Methods in Corrosion Science and Engineering*, CRC Taylor and Francis, Baton Roca (2005) 760 pp.

Bockris, J. O'M, Reddy A. K. N., *Modern Electrochemistry*, vol. 2 ,Plenum Press, New York (1970) pp. 623 – 1432.

Morrison, S. R., *Electrochemistry at Semiconductor and Oxidized Metal Electrodes*, Plenum Press, New York (1980).

© 2007, first edition, EPFL Press

References

[1] S. Brunauer, P.H. Emmet, E. Teller, *J. Am. Chem. Soc.* 60, 309 (1938).
[2] S. R. Bare, G. A. Somorjai, *Encyclopedia of Physical Science and Technology*, vol. 13, Academic Press, 1987, p. 551.
[3] D. Briggs, M. P. Seah, *Practical Surface Analysis*, John Wiley, Chichester 1983, p. 112.
[4] M. P. Seah, W. A. Dench, *Surf. Interf. Anal.* 1, 2 (1979).
[5] S. Mischler, A. Vogel, H. J. Mathieu, D. Landolt, *Corr. Science* 32, 925 (1991).
[6] P. Marcus, I. Olefjord, *Surf. Interf. Analysis* 11, 569 (1988).
[7] C. Feldman, J. W. Mayer, *Fundamentals of Surface and Thin Film Analysis*, North Holland, New York (1986) p. 267.
[8] C. Vickermann, *Chemistry in Britain*, Oct. 1987, p. 969.
[9] G. Binnig, H. Rohrer, *Helv. Chim. Acta*, 55, 726 (1982).
[10] P. Marcus, *Electrochim. Acta* 43, 109 (1998).
[11] G. Binnig, C. F. Quate, *Phys. Rev. Letters* 56, 930 (1986).
[12] Q. Wu, D. Barkey, *J. Electrochem. Soc.* 144, L261 (1997).
[13] Norm ISO No. 468.
[14] P.-F. Chauvy, C. Madore, D. Landolt, *Surf. Coat. Technol.* 110, 48 (1998).
[15] Bockris, J. O'M, Reddy A. K. N., *Modern Electrochemistry*, vol 2., Plenum Press, New York (1970), p 725.
[16] B. Conway, *Electrode Processes*, Ronald Press, New York (1965) p. 33.
[17] Parsons, R., *Aquatic Surface Chemistry*, W. Stumm, ed., J. Wiley, New York (1987) p. 33.
[18] P. Delahay, *Double Layer and Electrode Kinetics*, Wiley Interscience, New York (1965) p. 20.
[19] Reeves R., *Comprehensive Treatise of Electrochemistry,* vol. 1, J. O'M Bockris, B. E. Conway, E. Yeager, eds., Plenum Press, New York (1980) p. 87.
[20] Pleskov, Yu. V., *Comprehensive Treatise of Electrochemistry*, vol. 1, J. O'M Bockris, B. E. Conway, E. Yeager, eds., Plenum Press, New York (1980) p. 291.
[21] Reference 20, p. 316.
[22] E. McCafferty, *J. Electrochem. Soc.* 146, 2863 (1999).
[23] M. J. Kosmulski, *J. Colloid Interface Sci.* 275, 214 (2004).
[24] P. W. Schindler, W. Stumm, *Aquatic Surface Chemistry*, W. Stumm, ed., J. Wiley, New York (1987) p. 97, 101.

© 2007, first edition, EPFL Press

CHAPTER 4

CORROSION REACTION RATES

4.1 BASIC CONCEPTS

4.1.1 Single and mixed electrodes

Whenever only one electrode reaction takes place on a metal surface in a given solution, that system is called a ***single electrode***. This is the case for copper immersed into de-aerated and slightly acidic copper sulfate solution:

$$Cu^{2+} + 2e = Cu \tag{4.1}$$

The ***open circuit potential*** (ocp) is the potential set up spontaneously by an electrode in the absence of an external current. For a single electrode, the open circuit potential is equal to the equilibrium potential, E_{rev}.

If the potential of an electrode differs from the open circuit potential, an electric current is crossing the electrode-electrolyte interface. The ***overpotential*** η is defined as the difference between the electrode potential and the equilibrium potential of an electrode reaction:

$$\eta = E - E_{rev} \tag{4.2}$$

A positive overpotential indicates that an anodic current is crossing the interface; a negative one means that the current is cathodic.

Most often, several electrode reactions take place simultaneously at a metal-electrolyte interface. Such systems are referred to as ***mixed electrodes***. If, in the previous example, the copper sulfate solution is aerated, two electrode reactions (called *partial reactions*) are observed at the open circuit potential; the oxidation of copper

$$Cu \rightarrow Cu^{2+} + 2e \tag{4.3}$$

and the reduction of oxygen:

$$\frac{1}{2}O_2 + 2H^+ + 2e \rightarrow H_2O \tag{4.4}$$

© 2007, first edition, EPFL Press

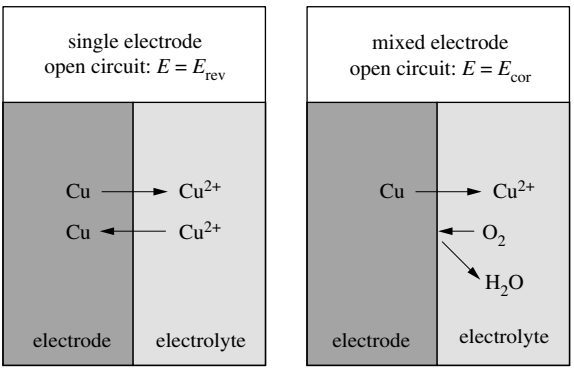


Figure 4.1 A copper electrode in contact with itw own ions (single electrode) and with an aerated solution (mixed electrode).

The corresponding overall reaction is

$$Cu + \frac{1}{2}O_2 + 2H^+ \rightarrow Cu^{2+} + H_2O \tag{4.5}$$

The copper thus corrodes without any external current. The open circuit potential of a mixed electrode undergoing corrosion, is called the ***corrosion potential*** (in the literature it is sometimes also called the *free corrosion potential*). The corrosion potential has a value that lies in between the equilibrium potentials of the partial electrode reactions. In contrast to the equilibrium potential, which is a thermodynamic quantity, the corrosion potential is determined by kinetics; its value depends on the rates of both the anodic and the cathodic partial reactions present.

The ***polarization*** ζ expresses the difference between the potential of a mixed electrode subjected to anodic or cathodic polarization and its corrosion potential.

$$\zeta = E - E_{cor} \tag{4.6}$$

A polarization of $\zeta > 0$ indicates an anodic and a polarization of $\zeta < 0$ a cathodic current flow. Figure 4.1 schematically shows the difference between single and mixed electrodes. Corrosion of metals always involves mixed electrodes.

4.1.2 Polarization curves

The rate of an electrode reaction depends on the potential drop at the electrode-electrolyte interface. According to Faraday's law (equation 1.8) the rate of reaction is proportional to the current density that flows through the electrode-electrolyte interface. By measuring the current density as a function of potential we can therefore get information about the kinetics of electrochemical reactions. The functional dependence between current density and potential is called ***polarization curve***. To experimentally determine a polarization curve one can control either the potential or the current and measure the other quantity. One thus obtains a ***potentiostatic*** polarization curve, $i = f(E)$, or a ***galvanostatic*** polarization curve, $E = f(i)$, respectively.

© 2007, first edition, EPFL Press

Potentiostatic mode

To measure a potentiostatic polarization curve one uses a *dc* current source called potentiostat and a three-electrode arrangement comprising the working electrode, the counter electrode and a reference electrode (Figure 4.2). A ***potentiostat*** is essentially an electronic amplifier that regulates the current between the working electrode and the counter electrode in such a way as to keep the potential difference between the working electrode and the reference electrode constant. To measure a potentiostatic polarization curve one imposes a constant potential and measures the corresponding current. The time to reach steady state depends on the specific conditions of the experiment. Alternatively, one can impose a linear potential scan and obtain a *potentiodynamic polarization curve*. Depending on the potential sweep rate and the characteristics of the electrochemical system such curves may or may not correspond to steady state current conditions.

To obtain the ***current density*** i (A m^{-2}) one divides the measured current I by the surface A of the working electrode:

$$i = I / A \tag{4.7}$$

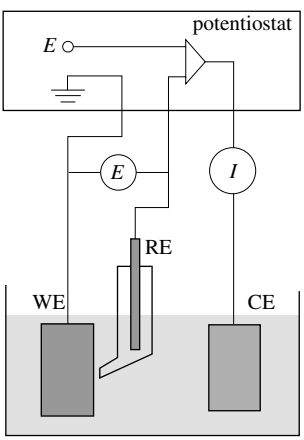


Figure 4.2 Experimental set up for potentiostatic polarization. WE: working electrode, RE: reference electrode, CE: counter electrode.

Galvanostatic mode

In order to produce a galvanostatic polarization curve, one uses a galvanostat as a *dc* current source and measures the potential difference between the working electrode and the reference electrode as a function of the applied current. A high-impedance voltmeter, called an electrometer (≥ 10 MΩ), prevents the current from passing through the reference electrode loop, which would introduce an error. A potentiostat can normally be used as a galvanostat, requiring only a change of connections (Figure 4.3).

© 2007, first edition, EPFL Press

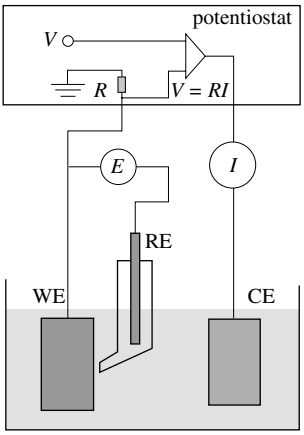


Figure 4.3 Experimental set up for galvanostatic polarization. WE: working electrode, RE: reference electrode, CE: counter electrode.

In principal, the paired values of potential and current density should not depend on the experimental method used. However, under certain conditions, the polarization curves acquired under potential control are not identical to those acquired under current control. In particular, when the polarization curve exhibits a maximum or a plateau region, more information is contained in a potentiostatic measurement (Figure 4.4). On the other hand, the galvanostatic technique is of interest when the slope di/dE is steep, and a slight variation of potential leads to a large change in the measured current. Modern potentiostats usually include software that permits to select the desired mode of operation and to acquire and plot experimental data.

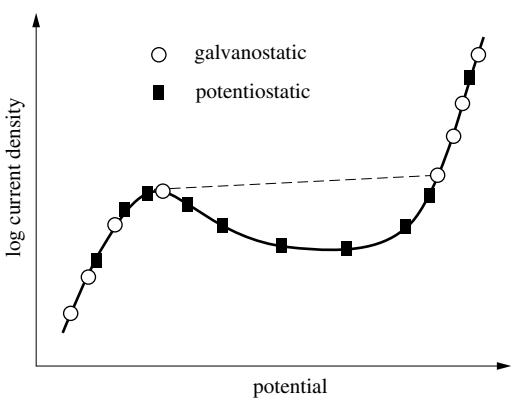


Figure 4.4 Polarization curve with a maximum plotted from potentiostatic and galvanostatic measurements.

© 2007, first edition, EPFL Press

Influence of the ohmic drop in the electrolyte

The current that passes between the working electrode and the counter electrode creates a potential gradient in the electrolyte. The effect is particularly pronounced at high current densities or in electrolyte solutions of low conductivity (Sect. 4.3). The measured potential, E_{measured}, is the sum of the electrode potential E and the ohmic potential drop in the electrolyte between the reference electrode and the working electrode, $\Delta\Phi_\Omega$:

$$E_{\text{measured}} = E + \Delta\Phi_\Omega \tag{4.8}$$

When measuring or controlling the potential of an electrode one wants to minimize the effect of $\Delta\Phi_\Omega$. The usual method calls for the placement of the reference electrode inside an electrolyte-filled recipient that is in contact with the cell via a thin capillary (***Luggin capillary***), whose extremity is placed close to the working electrode. As a first approximation, the value of the term $\Delta\Phi_\Omega$ can then be estimated from equation (4.9).

$$\Delta\Phi_\Omega = \frac{i \cdot L}{\kappa} \tag{4.9}$$

where L is the distance between the extremity of the capillary and the surface of the working electrode and κ represents the conductivity of the electrolyte.

To minimize $\Delta\Phi_\Omega$ the extremity of the Luggin capillary should be placed very close to the surface of the working electrode. However, if the distance L becomes too small, the current flow at the working electrode near the point of measurement will be perturbed due to shielding and this introduces again an error in the measurement. A good compromise is to select a distance of about three times the diameter d_c of the capillary (Figure 4.5).

$$L \approx 3d_c \tag{4.10}$$

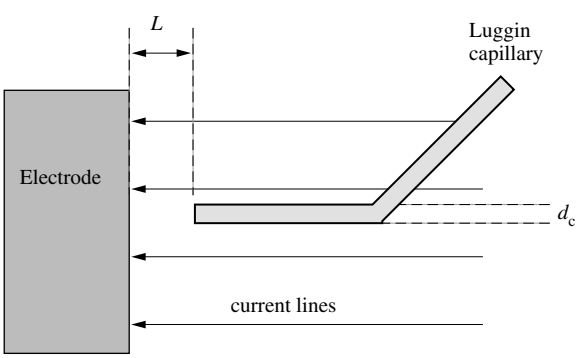


Figure 4.5 Optimal placement of Luggin capillary.

© 2007, first edition, EPFL Press

The ohmic potential drop $\Delta\Phi_\Omega$ often poses a practical limit to the precision of potential measurements. The use of an appropriate Luggin capillary helps to reduce its effect. In some cases one may mathematically correct for $\Delta\Phi_\Omega$ using equation 4.9 or using a measured value. In Section 5.2 experimental methods will be presented that permit to experimentally determine $\Delta\Phi_\Omega$ for any electrode arrangement.

4.1.3 Rate-limiting step of electrochemical reactions

Corrosion reactions consist of at least one anodic partial reaction and one cathodic partial reaction, each of which involves several steps (Figure 4.6). The overall reaction rate is limited by the rate of the slowest partial reaction. We can thus make a distinction between corrosion reactions under anodic or cathodic control.

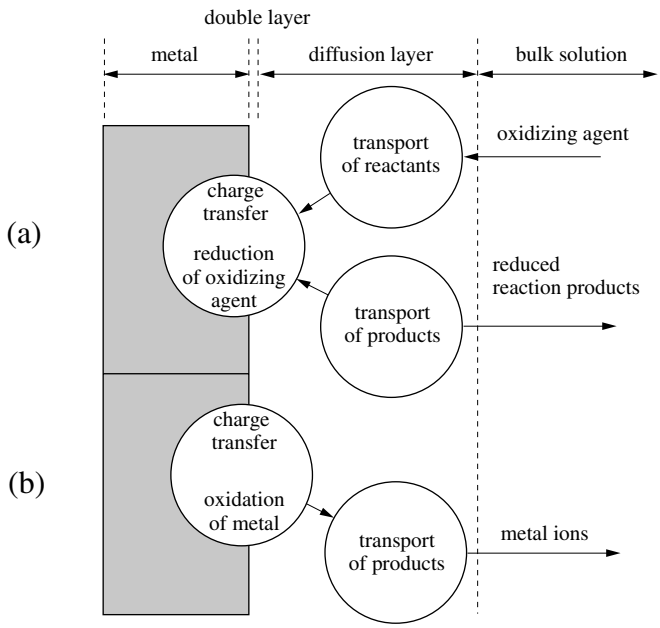


Figure 4.6 Reaction steps during the corrosion of a metal in liquid environments: (a) cathodic partial reaction, (b) anodic partial reaction.

Cathodic partial reactions

The oxidizing agent present in the electrolyte diffuses towards the electrode surface, where it reacts by accepting one or several electrons. In certain cases, the charge transfer reaction at the electrode takes place just before or after a heterogeneous chemical reaction taking place at the electrode-electrolyte interface or a homogeneous chemical reaction taking place in the electrolyte solution near the electrode surface. For example, the reduction of two protons can produce two adsorbed hydrogen atoms that subsequently, via a heterogeneous chemical reaction, form a hydrogen molecule

© 2007, first edition, EPFL Press

(Chapter 5). The hydrogen then leaves the electrode, either dissolved in solution or in the gaseous state.

Mass transport phenomena taking place before or after charge transfer determine the concentration of the reactants and products at the electrode surface. The electrolyte layer contiguous to the surface, in which the concentration of the reactants or products differs from that in the bulk electrolyte is called the ***diffusion layer***. The thickness of the diffusion layer depends on the prevailing convection conditions and is typically between 1 and 100 μm. This is much more than that of the electric double layer, which is typically on the order of 0.2–10 nm only.

Anodic partial reactions

The anodic partial reaction also involves a charge transfer at the interface because a metal atom loses electrons. It then dissolves in the solution as a hydrated or complexed ion and diffuses towards the bulk. In the vicinity of the metal surface, the concentration generated by dissolution therefore often exceeds that of the bulk electrolyte. Once the solubility threshold is reached, solid reaction products begin to precipitate and form a porous film. Alternatively, under certain conditions, metal ions do not dissolve at all but form a thin compact oxide layer, called ***passive film***. The properties of the passive film then determine the rate of corrosion of the underlying metal (Chapter 6).

Rate-limiting step of corrosion reactions

From a kinetic point of view, we can identify three types of corrosion reactions, depending on the rate-limiting step:

- Corrosion controlled by the kinetics of a charge-transfer reaction, either anodic or cathodic, at the metal-electrolyte interface. An example is the corrosion of steel in acids.
- Corrosion controlled by the rate of mass transport of the oxidizing agent or of anodic reaction products. A well-known example is the corrosion of steel in neutral, aerated solution.
- Corrosion controlled by the properties of the passive film. The reaction is then under anodic control and the average corrosion rate is often quite small (although locally it may be substantial, see Chapter 7). Stainless steels in aqueous solutions behave in this manner.

4.2 CHARGE TRANSFER

4.2.1 Butler-Volmer equation of a single electrode

This section describes the electrode reactions that are limited by the rate of charge transfer at the electrode-electrolyte interface. In this case, the ***Butler-Volmer equation*** (BV equation) provides a functional relationship between the potential and the current

© 2007, first edition, EPFL Press

density. A simple model, based on the absolute rate theory of chemical reactions, will serve to derive the BV equation. Let us consider the electrode reaction (4.11), which involves an electron transfer between a dissolved Fe^{2+} ion and an inert electrode, such as platinum, leading to formation of Fe^{3+}.

$$Fe^{2+} \underset{v_c}{\overset{v_a}{\rightleftarrows}} Fe^{3+} + e \tag{4.11}$$

According to Faraday's law, the current density at the electrode, i, is proportional to the rate of reaction v (mol m^{-2} s^{-1}). Since the reaction can take place in both directions, the net rate corresponds to the difference between the rate of oxidation of Fe^{2+}, v_a , and the rate of reduction of Fe^{3+}, v_c. The subscripts a and c refer to anodic and cathodic processes, respectively. With the charge number $n = 1$ we obtain (4.12).

$$i = F\,v = F\,(v_a - v_c) \tag{4.12}$$

The rates v_a and v_c are, respectively, proportional to the concentrations of Fe^{2+} and Fe^{3+} at the electrode surface and they obey Arrhenius' law.

$$v_a = k'_a \cdot c_{Fe^{2+},s} \exp\left(-\frac{\Delta G_a^{\#}}{RT}\right) \tag{4.13}$$

$$v_c = k'_c \cdot c_{Fe^{3+},s} \exp\left(-\frac{\Delta G_c^{\#}}{RT}\right) \tag{4.14}$$

Here, k'_a and k'_c are constants; $c_{Fe2+,s}$ and $c_{Fe3+,s}$ are the concentrations of Fe^{2+} and Fe^{3+} at the electrode-electrolyte interface (outside the double layer); and $\Delta G_a^{\#}$ and $\Delta G_c^{\#}$ represent the Gibbs free energies of activation of the anodic and cathodic partial reactions. According to absolute rate theory

$$k'_a = k'_c = v\frac{kT}{h} \tag{4.15}$$

where v refers to the transition probability ($0 < v < 1$). For an electrochemical reaction, the Gibbs free energy of activation depends on the potential difference across the double layer $\Delta\Phi$. By making the assumption that $\Delta G_a^{\#}$ and $\Delta G_c^{\#}$ vary linearly with $\Delta\Phi$, one can express the free energies of activation as follows:

$$\Delta G_a^{\#} = \Delta G_{a,ch}^{\#} - \alpha F \Delta\Phi \tag{4.16}$$

$$\Delta G_c^{\#} = \Delta G_{c,ch}^{\#} + (1-\alpha) F \Delta\Phi \tag{4.17}$$

Here, $\Delta G_{a,ch}^{\#}$ and $\Delta G_{c,ch}^{\#}$represent the "chemical" part of the Gibbs free energy of activation, which does not depend on the applied potential. The proportionality constant α is called the ***charge transfer coefficient***. Its value is situated between $0 < \alpha < 1$.

© 2007, first edition, EPFL Press

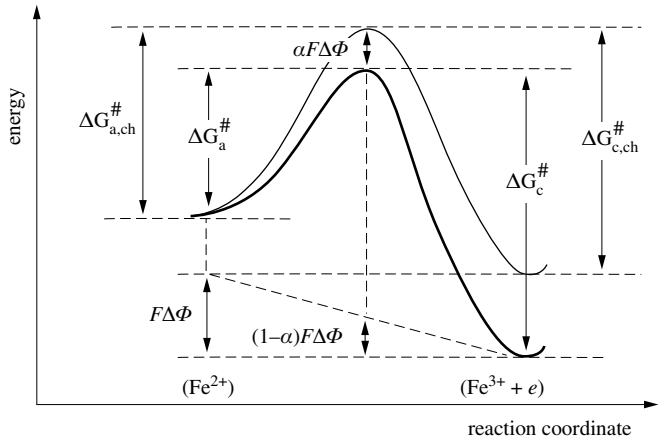


Figure 4.7 Influence of the potential on the activation energy for charge transfer at electrode-electrolyte interface.

Figure 4.7 schematically shows the influence of the potential difference $\Delta\Phi$ on the activation energy of the process leading from the reduced state (Fe^{2+}) to the oxidized state (Fe^{3+} + e), and for the reverse process. For $\Delta\Phi = 0$, the activation energy corresponds to $\Delta G^{\#}_{a,ch}$ and $\Delta G^{\#}_{c,ch,}$ respectively. If a positive potential is applied to the electrode, ($\Delta\Phi > 0$), the electron energy in the metal is lowered by the amount $F\ \Delta\Phi$. The metal therefore more easily accepts electrons, and the activation energy for the oxidation of the Fe^{2+} ion decreases, while that for the reverse reaction increases.

As long as the curve shape of the activation energy does not vary with potential, the charge-transfer coefficient is a constant. Figure 4.7 shows the case for the value $\alpha = 0.5$, resulting when the curve's maximum is located at the mid-point between the initial and final states. For this reason, α is sometimes referred to as the *symmetry factor.*

In order to obtain the Butler-Volmer equation, we group together all potential independent terms into the constants k''_a and k''_c.

$$k''_a = k'_a \exp\left(-\frac{\Delta G^{\#}_{a,ch}}{RT}\right) \tag{4.18}$$

$$k''_c = k'_c \exp\left(-\frac{\Delta G^{\#}_{c,ch}}{RT}\right) \tag{4.19}$$

With (4.12) and (4.13-14), this gives

$$v = k''_a\, c_{Fe2+,s} \exp\left(\frac{\alpha F}{RT}\Delta\Phi\right) - k''_c\, c_{Fe3+,s} \exp\left(-\frac{(1-\alpha)F}{RT}\Delta\Phi\right) \tag{4.20}$$

© 2007, first edition, EPFL Press

The potential of the working electrode E measured with respect to a given reference electrode is given by (4.21).

$$E = \Delta\Phi + \text{constant} \tag{4.21}$$

By introducing (4.21) into (4.20), and then grouping together the constant terms into k_a and k_c, we obtain:

$$v = k_a c_{Fe^{2+},s} \exp\left(\frac{\alpha F}{RT} E\right) - k_c c_{Fe^{3+},s} \exp\left(-\frac{(1-\alpha)F}{RT} E\right) \tag{4.22}$$

The values of k_a and k_c vary as a function of the selected reference electrode, because E also depends on this parameter. Normally, one indicates the potentials (and therefore also the corresponding rate constants) with respect to the standard hydrogen electrode.

The current density, i, is the sum of the partial anodic current density i_a and the partial cathodic current density, i_c:

$$i = i_a + i_c \tag{4.23}$$

By convention, the anodic current density is positive ($i_a > 0$), and the cathodic current density is negative ($i_c < 0$). The expressions (4.12), (4.22) and (4.23) yield (4.24), the Butler-Volmer equation for the half-cell reaction $Fe^{3+} + e = Fe^{2+}$.

$$i = i_a + i_c = F k_a c_{Fe^{2+},s} \exp\left(\frac{\alpha F}{RT} E\right) - F k_c c_{Fe^{3+},s} \exp\left(-\frac{(1-\alpha)F}{RT} E\right) \tag{4.24}$$

For a given electrode reaction, involving the transfer of n electrons,

$$B_{ox} + n e = B_{red} \tag{4.25}$$

we find by analogous reasoning:

$$i = i_a + i_c = n F k_a c_{red,s} \exp\left(\frac{\alpha n F}{RT} E\right) - n F k_c c_{ox,s} \exp\left(-\frac{(1-\alpha) n F}{RT} E\right) \tag{4.26}$$

In this equation, $c_{ox,s}$ and $c_{red,s}$ represent, respectively, the concentrations of B_{ox} et B_{red} at the electrode surface. The expression (4.26) represents the Butler-Volmer equation for any first-order electrode reaction. Its rate, by definition, is proportional to the reactant concentration at the electrode surface.

Exchange current density

At equilibrium, the net reaction rate is zero. This does not mean, however, that the partial reactions have stopped:

$$E = E_{rev} : \qquad i = i_a + i_c = 0. \tag{4.27}$$

© 2007, first edition, EPFL Press

As a result:

$$i_{a(E_{rev})} = -i_{c(E_{rev})} = i_0 \tag{4.28}$$

The equation (4.28) defines the ***exchange current density*** i_0 of an electrode reaction. This new quantity characterizes the rate of charge transfer at equilibrium.

Under equilibrium conditions, the species B_{ox} and B_{red} involved in the charge-transfer reaction each have the same concentration at the electrode surface (subscript *s*) as in the bulk solution (subscript *b*):

$$c_{red,s} = c_{red,b} \tag{4.29}$$

$$c_{ox,s} = c_{ox,b} \tag{4.30}$$

The relations (4.28) to (4.30) then give:

$$i_0 = n\,F\,k_a\,c_{red,b}\exp\left(\frac{\alpha n F}{RT}E_{rev}\right) = n\,F\,k_c\,c_{ox,b}\exp\left(-\frac{(1-\alpha)nF}{RT}E_{rev}\right) \tag{4.31}$$

When this equation is plugged into (4.26) yet another form of the Butler-Volmer equation is obtained:

$$i = i_0\frac{c_{red,s}}{c_{red,b}}\exp\left(\frac{\alpha n F}{RT}\left(E-E_{rev}\right)\right) - i_0\frac{c_{ox,s}}{c_{ox,b}}\exp\left(-\frac{(1-\alpha)nF}{RT}\left(E-E_{rev}\right)\right) \tag{4.32}$$

Tafel coefficients

We can now bring the overvoltage $\eta = E - E_{rev}$ explicitly into (4.32) and define the anodic ***Tafel coefficient***

$$\beta_a = \frac{RT}{\alpha\,n\,F} \tag{4.33}$$

and the cathodic Tafel coefficient,

$$\beta_c = \frac{RT}{(1-\alpha)nF} \tag{4.34}$$

which then gives

$$i = i_0\frac{c_{red,s}}{c_{red,b}}\exp\left(\frac{\eta}{\beta_a}\right) - i_0\frac{c_{ox,s}}{c_{ox,b}}\exp\left(-\frac{\eta}{\beta_c}\right) \tag{4.35}$$

If the reactant and product concentrations are uniform in the electrolyte ($c_{red,s} = c_{red,b}$, $c_{ox,s} = c_{ox,b}$) the expression becomes:

$$i = i_0\exp\left(\frac{\eta}{\beta_a}\right) - i_0\exp\left(-\frac{\eta}{\beta_c}\right) \tag{4.36}$$

© 2007, first edition, EPFL Press

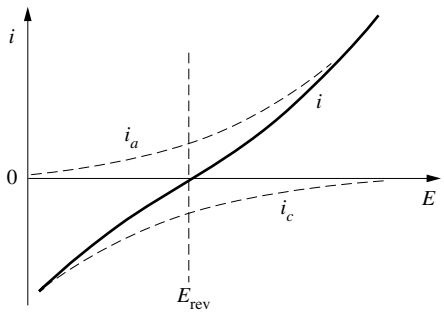


Figure 4.8 Partial current densities and total current density near the equilibrium potential.

This is the most common form of the Butler-Volmer equation. It is applicable to electrode reactions whose rate is entirely limited by charge transfer at the interface. This process is sometimes called ***activation control***; the corresponding overpotential is then called the ***activation overpotential***.

Figure 4.8 provides a graph that shows, on a linear scale, the current density i as a function of the potential according to (4.36), as well as the anodic and cathodic partial current densities. According to (4.27), the current density at the equilibrium potential is zero.

The values of the Tafel coefficients β_a and β_c depend on the mechanism of the electrode reactions, which often consist of several elementary steps (Section 5.1). It is however not necessary to know the mechanism in order to apply the Butler-Volmer equation. Indeed, equation (4.36) describes the charge-transfer kinetics in a global, mechanism-independent fashion, making reference to three easily measured quantities: i_0, β_a and β_c. The formulae (4.37) and (4.38) then define the anodic and cathodic Tafel coefficients:

$$\beta_a = \frac{dE}{d\ln i_a} \tag{4.37}$$

$$\beta_c = -\frac{dE}{d\ln|i_c|} \tag{4.38}$$

Tafel lines

In order to determine experimentally the kinetic parameters i_0, β_a and β_c, it is best to present the current density on a logarithmic scale, because this highlights the linear relationship between the logarithm of i and the overpotential found whenever this latter parameter, in absolute value, is large.

The region of potential that corresponds to $\eta/\beta_a >> 1$ is called the ***anodic Tafel region***. Equation (4.36) then becomes:

© 2007, first edition, EPFL Press

$$i = i_a = i_0 \exp\left(\frac{\eta}{\beta_a}\right) \tag{4.39}$$

and taking the logarithm yields:

$$\eta = -\beta_a \ln i_0 + \beta_a \ln i \tag{4.40}$$

By converting to base-10 logarithms and defining the anodic Tafel constants a_a and b_a , we obtain the Tafel equation of a simple anodic reaction, also called the *anodic Tafel line*:

$$\eta = a_a + b_a \log i \tag{4.41}$$

$$a_a = -2.303\, \beta_a \log i_0$$

$$b_a = 2.303\, \beta_a$$

In an analogous manner, for the ***cathodic Tafel region***, $\eta/\beta_c << -1$:

$$i = i_c = -i_0 \exp\left(-\frac{\eta}{\beta_c}\right) \tag{4.42}$$

$$\eta = \beta_c \ln i_0 - \beta_c \ln |i| \tag{4.43}$$

With these, it is possible to write expressions for the cathodic Tafel constants a_c and b_c of the Tafel equation of a cathodic reaction:

$$\eta = a_c - b_c \log |i| \tag{4.44}$$

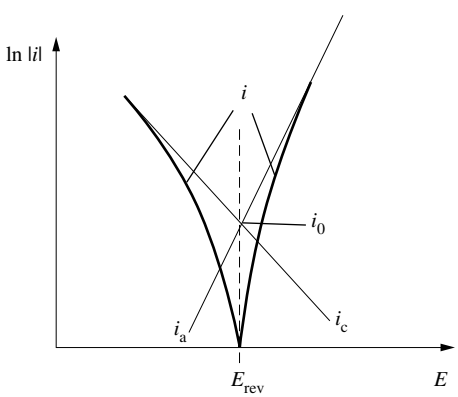


Figure 4.9 Evans diagram showing anodic and cathodic partial current densities and total current density of an electrode reaction on a logarithmic scale.

© 2007, first edition, EPFL Press

$$a_c = 2.303\,\beta_c \log i_0$$

$$b_c = 2.303\,\beta_c$$

The ***Tafel equations*** (4.41) and (4.44) describe the anodic and cathodic limits of the Butler-Volmer equation. Their discovery by Tafel at the beginning of the twentieth century took place many years before the formulation of the rate equation by Butler and Volmer.

The extrapolation of an experimental polarization curve, measured from the Tafel region to the reversible potential, reveals the exchange current density i_0 (Figure 4.9). The reciprocal of the slopes of the straight lines yield the Tafel coefficients β_a and β_c.

Value of the exchange current density and of the Tafel coefficients

For different electrode reactions, the exchange current density can vary over several orders of magnitude (10^{-12} to 1 A cm^{-2}). Even for a given electrode reaction, its value may differ greatly depending on electrolyte composition or the state of the metal surface. The value of i_0 for a given system is therefore generally not known *a priori* and must be measured

Figure 4.10 and Figure 4.11 show the influence of i_0 on the shape of the polarization curve. The curves were calculated with the assumption that $\alpha = 0.5$ and $n = 2$. On a linear scale, they become flatter close to the reversible potential as i_0

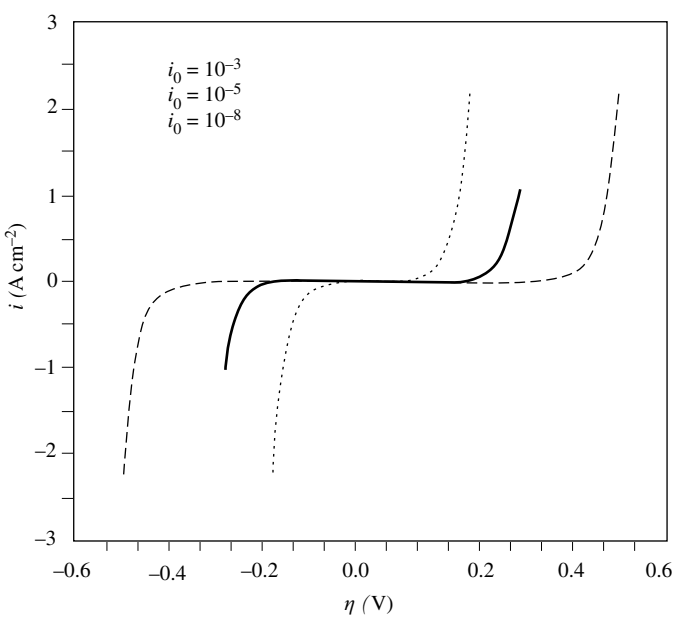


Figure 4.10 Polarization curves (linear scale) calculated using the Butler-Volmer equation, with $n = 2$ and $\alpha = 0.5$, for different exchange current densities.

© 2007, first edition, EPFL Press

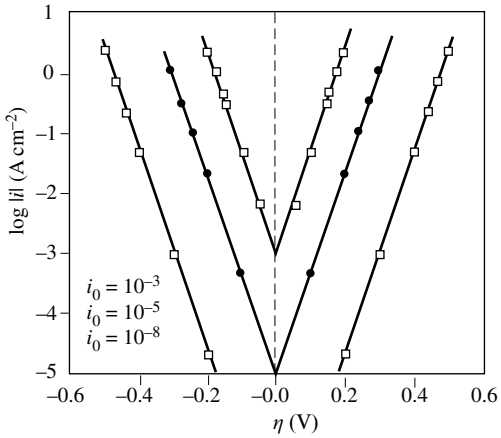


Figure 4.11 Polarization curves (logarithmic scale) calculated using the Butler-Volmer equation, with $n = 2$ and $\alpha = 0.5$, for different exchange current densities.

decreases. If i_0 approaches zero, the current density in this region becomes very weak, corresponding to the behavior of an ***ideally polarizable*** electrode, characterized by a high charge-transfer resistance (Sect. 3.5). At the other extreme, when i_0 becomes large, the polarization curve no longer shows a point of inflection at the reversible potential. The electrode is then referred to as ***non-polarizable***. This means that the overpotential remains small even at high current density, and the electrode potential hardly deviates from the equilibrium potential ($E \approx E_{rev}$).

In contrast to the exchange current density, the Tafel coefficients β_a and β_c vary relatively little between different reactions: their values are typically in the range of 20–50 mV.

4.2.2 Butler-Volmer equation of a mixed electrode

The theory of mixed electrodes was developed by Wagner and Traud [1] and first applied to corrosion by Stern and Geary [2]. In the following we develop the Butler-Volmer equation for a mixed electrode by taking a specific example: the uniform corrosion of iron in hydrochloric acid in the absence of concentration gradients ($c_{i,s} = c_{i,b}$). Two electrode reactions take place simultaneously:

$$\mathrm{Fe} \rightarrow \mathrm{Fe}^{2+} + 2e \tag{4.45}$$

$$2\,\mathrm{H}^+ + 2e \rightarrow \mathrm{H}_2 \tag{4.46}$$

$$\mathrm{Fe} + 2\,\mathrm{H}^+ \rightarrow \mathrm{Fe}^{2+} + 2\,\mathrm{H}_2 \tag{4.47}$$

© 2007, first edition, EPFL Press

If we connect iron as the working electrode in an electrochemical cell, and then apply an anodic or cathodic potential, a current flows. Its value divided by the electrode surface A corresponds to the sum of all partial current densities.

$$i = I / A = i_{Fe} + i_H = i_{a,Fe} + i_{c,Fe} + i_{a,H} + i_{c,H} \tag{4.48}$$

The subscripts Fe and H refer to the respective electrode reactions. Near the corrosion potential, the contributions of $i_{c,Fe}$ and $i_{a,H}$ are generally negligible ($i_{c,Fe} \approx i_{a,H} \approx 0$), which allows us to simplify Equation (4.48):

$$i = i_{a,Fe} + i_{c,H} \tag{4.49}$$

At the corrosion potential ($E = E_{cor}$), the current density i is zero.

$$i = i_{a,Fe} + i_{c,H} = 0 \tag{4.50}$$

We assume, in the first place, that the charge transfer at the metal-solution interface represents the rate-limiting step and, secondly, that the kinetics of the two partial reactions are independent of one another. The partial current density of each reaction then obeys the Butler-Volmer reaction.

$$i_{Fe} = i_{a,Fe} + i_{c,Fe} = i_{0,Fe} \exp\left(\frac{\eta_{Fe}}{\beta_{a,Fe}}\right) - i_{0,Fe} \exp\left(-\frac{\eta_{Fe}}{\beta_{c,Fe}}\right) \tag{4.51}$$

$$i_H = i_{a,H} + i_{c,H} = i_{0,H} \exp\left(\frac{\eta_H}{\beta_{a,H}}\right) - i_{0,H} \exp\left(-\frac{\eta_H}{\beta_{c,H}}\right) \tag{4.52}$$

In these equations, the overpotentials η_{Fe} and η_H are defined as:

$$\eta_{Fe} = E - E_{rev,Fe} \tag{4.53}$$

$$\eta_H = E - E_{rev,H} \tag{4.54}$$

Combining equations (4.50), (4.51) and (4.52) yields equation (4.55), which defines the ***corrosion current density*** i_{cor}.

$$i_{0,Fe} \exp\left(\frac{E_{cor} - E_{rev,Fe}}{\beta_{a,Fe}}\right) = i_{0,H} \exp\left(-\frac{E_{cor} - E_{rev,H}}{\beta_{c,H}}\right) = i_{cor} \tag{4.55}$$

The corrosion current density is equal to the anodic partial current density at the corrosion potential. Its value, and therefore the rate of corrosion, depends on the kinetic parameters of both electrode reactions involved in the corrosion process.

$$i_{cor} = i_{a,Fe(E_{cor})} = -i_{c,H(E_{cor})} \tag{4.56}$$

© 2007, first edition, EPFL Press

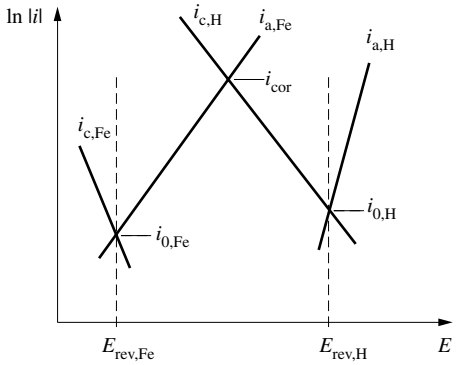


Figure 4.12 Evans diagram of the corrosion of iron in acidic environment.

At the corrosion potential, the rate of corrosion v_{cor} of iron is thus equal to:

$$v_{cor} = \frac{i_{a,Fe}}{2F} = \frac{i_{cor}}{2F} \tag{4.57}$$

In order to derive the Butler-Volmer equation for the mixed electrode, we replace the overpotentials that appear in equations (4.51) and (4.52) by the polarization $\zeta = E - E_{cor}$:

$$\eta_{Fe} = E - E_{rev,Fe} = \zeta + (E_{cor} - E_{rev,Fe}) \tag{4.58}$$

$$\eta_{H} = E - E_{rev,H} = \zeta + (E_{cor} - E_{rev,H}) \tag{4.59}$$

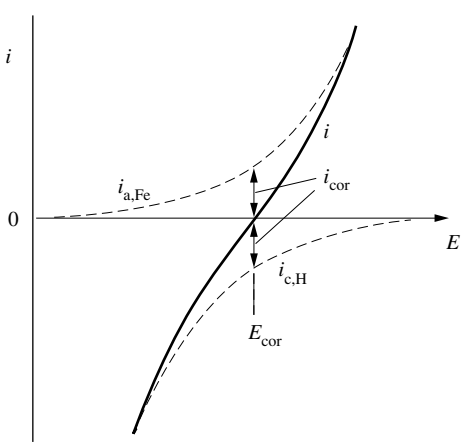


Figure 4.13 Partial current densities and total current density of mixed electrode near the corrosion potential.

© 2007, first edition, EPFL Press

With (4.55) we thus obtain:

$$i = i_{\mathrm{a,Fe}} + i_{\mathrm{c,H}} = i_{\mathrm{cor}} \exp\left(\frac{\zeta}{\beta_{\mathrm{a,Fe}}}\right) - i_{\mathrm{cor}} \exp\left(-\frac{\zeta}{\beta_{\mathrm{c,H}}}\right) \tag{4.60}$$

Figure 4.12 shows how the logarithm of the partial current densities varies as a function of the potential. This type of graph, describing the behavior of a mixed electrode, is called an ***Evans diagram***. The figure also shows E_{cor} and i_{cor}, as well as the reversible potentials of the two electrode reactions involved. In Figure 4.13, the variation of the current as a function of the potential, near the corrosion potential, is presented in linear scale. It permits to appreciate the relative importance of the anodic and cathodic partial current densities.

Generalization

The above reasoning is easily extended to a system of k electrode reactions, for example to the case of an alloy corroding in the presence of protons and oxygen. The measured current I is the sum of all anodic and cathodic partial currents:

$$I = \sum I_{\mathrm{a,k}} + \sum I_{\mathrm{c,k}} \tag{4.61}$$

At the corrosion potential:

$$I = \sum I_{\mathrm{a,k}(E_{\mathrm{cor}})} + \sum I_{\mathrm{c,k}(E_{\mathrm{cor}})} = 0 \tag{4.62}$$

Therefore:

$$I_{\mathrm{cor}} = \sum I_{\mathrm{a,k}(E_{\mathrm{cor}})} = -\sum I_{\mathrm{c,k}(E_{\mathrm{cor}})} \tag{4.63}$$

The rate of uniform corrosion at the corrosion potential is

$$v_{\mathrm{cor}} = \sum v_{\mathrm{cor,k}} = \sum \frac{I_{\mathrm{a,k}}}{n_{\mathrm{k}} F A} \qquad (\mathrm{mol/m^2 s}) \tag{4.64}$$

where A is the surface area of the corroding metal. An analogous development as described above for the specific case of corrosion of iron in acid solution yields the general form of the Butler-Volmer equation of a mixed electrode:

$$I = \sum I_{\mathrm{a,k}} + \sum I_{\mathrm{c,k}} = \sum I_{\mathrm{cor}} \exp\left(\frac{\zeta}{\beta_{\mathrm{a,k}}}\right) - \sum I_{\mathrm{cor}} \exp\left(-\frac{\zeta}{\beta_{\mathrm{c,k}}}\right) \tag{4.65}$$

Since equations (4.61–4.65) contain the current I rather than the current density $i = I/A$ they are valid even if the surface areas available for the anodic and cathodic partial reactions differ.

© 2007, first edition, EPFL Press

4.2.3 Experimental determination of corrosion rates

Two types of experimental tests are commonly used to determine the corrosion rate in solution:

- Immersion tests;
- Electrochemical tests.

Immersion tests

In immersion testing one usually determines the mass loss of a sample exposed to a corrosive environment for a certain time. Different experimental methods can be used to measure the mass of corroded metal: gravimetry, solution analysis, or in case of thin foil samples, the diminution of ohmic resistance. Volumetric gas analysis is an other method to monitor the rate of corrosion; for example by measuring the volume of oxygen that reacted with the metal or the volume of hydrogen produced by the corrosion reaction.

Figure 4.14 shows an experimental set-up often employed in immersion tests. Several samples are placed onto a non-metallic mount and then immersed into a corrosive solution, held at constant temperature. A condenser is used to prevent loss of liquid by evaporation. During the test, the mass loss of the samples is measured after defined periods of exposure by weighing after having cleaned the sample surface of all solid corrosion products. Different pickling solutions serving this purpose are listed in the literature [3,4].

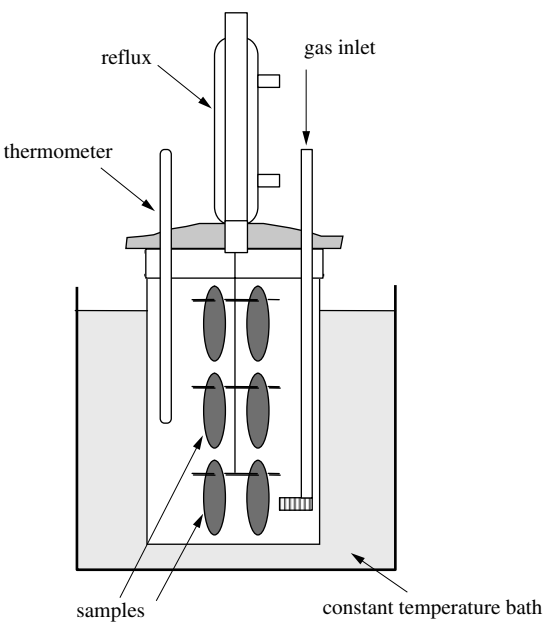


Figure 4.14 Experimental apparatus for corrosion testing by immersion.

© 2007, first edition, EPFL Press

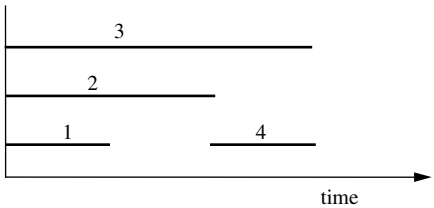


Figure 4.15 Planned interval testing.

Sometimes, changes with time of the rate of corrosion call for a procedure with mass loss measurements at several intervals. Figure 4.15 shows the schedule for a planned interval test with four identical samples. Samples 1 to 3 are used to determine the evolution with time of the rate of corrosion. This provides information about a number of parameters, including the role of surface films formed during corrosion. On the other hand, a comparison of the corrosion rate of samples 1 and 4 reveals possible changes in the corrosivity of the solution during the test.

Electrochemical corrosion tests

Immersion tests provide no information about reaction mechanisms and often they require relatively long exposure times. Electrochemical tests do not have these drawbacks and they are therefore widely used in practice. In the following electrochemical polarization methods are presented that provide information on the rate of uniform corrosion under conditions where the rate is controlled by charge-transfer. Other electrochemical test methods will be presented in subsequent chapters.

Extrapolation of Tafel lines

For a system comprised of a single anodic and single cathodic reaction of the type

$$M + \nu_{ox} B_{ox} \rightarrow M^{n+} + \nu_{red} B_{red} \tag{4.66}$$

where M refers to a metal and B_{ox} is an oxidizing agent, the equation (4.65) reads (by neglecting $I_{c,M}$ and $I_{a,B}$):

$$I = I_{cor} \exp\left(\frac{\zeta}{\beta_{a,M}}\right) - I_{cor} \exp\left(-\frac{\zeta}{\beta_{c,B}}\right) \tag{4.67}$$

At a polarizations exceeding 50 to 100 mV in the anodic or cathodic direction, one of the two exponential terms can be neglected, and the measured current is equal to the anodic or cathodic partial current, respectively. The extrapolation of the Tafel line to the corrosion potential then provides the value of I_{cor}. For example, for the cathodic Tafel region:

© 2007, first edition, EPFL Press

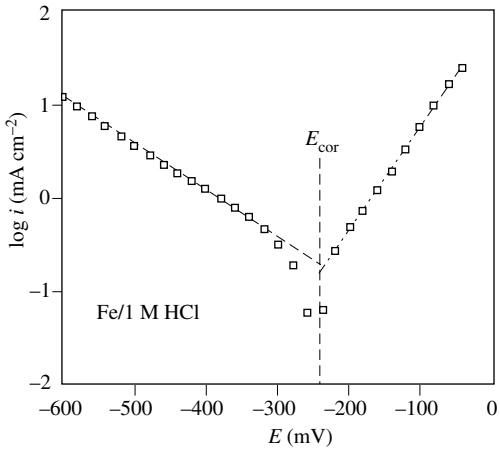


Figure 4.16 Logarithmic polarization curve of iron in 1M HCl; evaluation of i_{cor} by extrapolation of Tafel lines.

$$I = -I_{cor} \exp\left(-\frac{\zeta}{\beta_{c,B}}\right) \tag{4.68}$$

At the corrosion potential ($\zeta = 0$) equation (4.68) yields $I = I_{cor}$ and the rate of corrosion is

$$v_{cor} = \frac{I_{cor}}{n\,FA} \tag{4.69}$$

Figure 4.16 illustrates the described extrapolation. The same procedure can in principle be applied to the anodic Tafel region. However, because cathodic polarization does not corrode the sample, it is generally preferred.

We have seen that Tafel line extrapolation conveniently yields I_{cor} and by application of Faraday's law, the rate of corrosion at open circuit. Nevertheless, the method is based on certain assumptions that must be verified:

- the metal corrosion is uniform;
- the corrosion products are in the form of dissolved ions;
- the reaction mechanism does not change with applied potential;
- mass-transport effects are absent;
- ohmic effects ($\Delta\Phi_\Omega$) are negligible.

The first two conditions are fairly obvious and they are an inherent part of the described theory. The other conditions often limit the practical application of the method.

Firstly, because the electrode reaction mechanism greatly influences the Tafel slopes (Chap. 5) the reaction mechanism in the Tafel region should not differ from that prevailing at the corrosion potential. Mass transfer limitations and ohmic drops

© 2007, first edition, EPFL Press

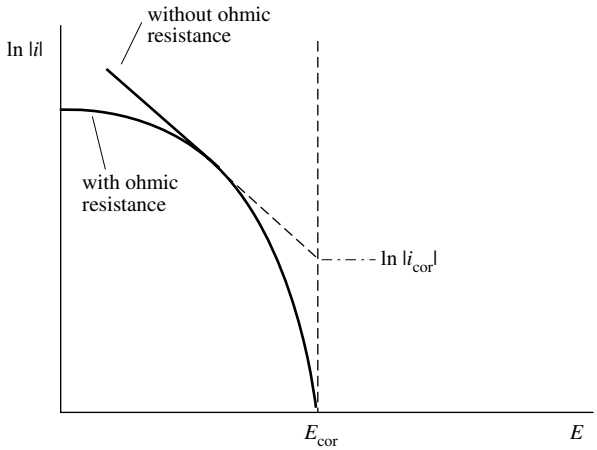


Figure 4.17 Effect of ohmic resistance between the working and reference electrodes on the shape of the measured polarization curves.

become more important at higher current density. The Tafel line extrapolation must be limited to conditions where the current density is well below the limiting current density (Sect. 4.3) and to solutions of sufficient conductivity to keep ohmic effects small. Figure 4.17 shows the effect of ohmic resistance on the shape of the cathodic polarization curve.

Tafel line extrapolation for the determination of corrosion rates is used mainly for laboratory testing conducted in acid electrolytes, a typical example being the evaluation of corrosion inhibitors.

Determination of polarization resistance

The polarization resistance method permits one to avoid certain difficulties associated with Tafel line extrapolation. The polarization resistance of an electrode, $r_{(E)}$, is defined by equation (4.70).

$$r(E) = \left.\frac{\mathrm{d}E}{\mathrm{d}i}\right|_E \qquad \left(\Omega\,\mathrm{m}^2\right) \tag{4.70}$$

Here, $r(E)$ depends on the potential because the Butler-Volmer relation is not linear. The ***polarization resistance at the corrosion potential***, r_{cor}, is equal to:

$$r_{\mathrm{cor}} = \left.\frac{\mathrm{d}E}{\mathrm{d}i}\right|_{E_{\mathrm{cor}}} = \left.\frac{\mathrm{d}\zeta}{\mathrm{d}i}\right|_{E_{\mathrm{cor}}} \qquad \left(\Omega\,\mathrm{m}^2\right) \tag{4.71}$$

To experimentally determine the polarization resistance at the corrosion potential one applies a polarization of a few millivolts in the cathodic and the anodic direction and determines the slope as schematically shown in Figure 4.18. Under these

© 2007, first edition, EPFL Press

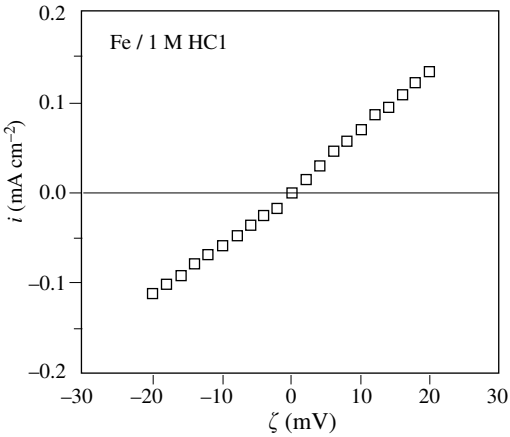


Figure 4.18 Polarization curve of Fe as measured in 1M HCl near the corrosion potential.

conditions the current density remains small and therefore the effects of the ohmic drop in the solution and of mass transport are less critical.

For a corrosion reaction controlled by charge transfer, differentiation of (4.67) at the corrosion potential ($\zeta = 0$) yields, after division by the surface A,

$$\frac{1}{A}\frac{\mathrm{d}I}{\mathrm{d}\zeta} = \frac{\mathrm{d}i}{\mathrm{d}\zeta} = \frac{i_{\mathrm{cor}}}{\beta_{\mathrm{a,M}}}\exp\left(\frac{\zeta}{\beta_{a,M}}\right) + \frac{i_{\mathrm{cor}}}{\beta_{\mathrm{c,B}}}\exp\left(-\frac{\zeta}{\beta_{\mathrm{c,B}}}\right) \tag{4.72}$$

$$r_{\mathrm{cor}} = \frac{1}{i_{\mathrm{cor}}}\frac{\beta_{\mathrm{a,M}}\cdot\beta_{\mathrm{c,B}}}{\beta_{\mathrm{a,M}}+\beta_{\mathrm{c,B}}} \tag{4.73}$$

Thus, if the Tafel coefficients are known, the measurement of the polarization resistance at the corrosion potential allows us to determine i_{cor}.

In practice, the exact values of $\beta_{\mathrm{a,M}}$ and $\beta_{\mathrm{c,B}}$ are often not needed, particularly if one is more interested in the variation of corrosion rate with time than in its precise magnitude. Indeed, Tafel coefficients for different electrode reactions lie typically in the range of 20–50 mV. As a consequence, when estimating the corrosion rate from polarization resistance the uncertainty in the rate of corrosion will hardly exceed a factor of two, even if one does not know the actual values of $\beta_{\mathrm{a,M}}$ and $\beta_{\mathrm{c,B}}$.

Effect of ohmic drop

In order to obtain a correct value for i_{cor} from the polarization resistance, it is necessary, on one hand, to measure under steady-state conditions (current stable with time) and, on the other hand, to avoid ohmic drop effects. Indeed, in presence of an ohmic drop between the reference and working electrodes, instead of r_{cor}, one measures an apparent polarization resistance $r_{\mathrm{ap,}}$ given by:

© 2007, first edition, EPFL Press

$$k''_{\mathrm{a}} = k'_{\mathrm{a}} \exp\left(-\frac{\Delta G^{\#}_{\mathrm{a,ch}}}{RT}\right) \tag{4.18}$$

$$r_{\mathrm{ap}} = \frac{\mathrm{d}}{\mathrm{d}i}\left(E + \Delta\Phi_{\Omega}\right) = r_{\mathrm{cor}} + r_{\Omega} \tag{4.74}$$

Here $r_{\Omega} = \mathrm{d}\Delta\Phi_{\Omega}/\mathrm{d}i$ represents the ohmic resistance in the electrolyte ($\Omega\ \mathrm{m}^2$).

Ohmic drops become particularly problematic for measurements carried out in weakly conducting electrolytes, for example, in potable water or in the soil. In this case, a technique with current interruption is sometimes used to experimentally determine r_{Ω}. When the current is cut, the ohmic contribution to the measured potential disappears instantaneously, whereas the potential across the double layer decays more slowly, analogous to a discharging capacitor (Chapter 5).

4.3 MASS TRANSPORT

4.3.1 Concentration overpotential

Often, the concentration of reactants and products in the bulk electrolyte differs from that found at the electrode surface. This leads to an overpotential that varies with the rate of transport of the reactants and products to and from the electrode.

Flux of minor ionic species

A ***minor ionic species*** is an ion that is present at low concentration in an electrolyte containing other ions at much higher concentration. If these other ions do not participate in the electrode reaction, they constitute the ***supporting electrolyte***. In a solution of 10^{-3} M $CuSO_4$ + 1 M H_2SO_4 for example, the Cu^{2+} ion is a minor ionic species, and sulfuric acid is the supporting electrolyte.

The potential gradient that exists in the elextrolyte contributes negligibly to the movement of minor ionic species; their transport is almost entirely by convection and diffusion. Therefore, the equations that are used for neutral species, such as for dissolved molecular oxygen, are also valid for minor ionic species. ***Convection*** refers to the macroscopic movement of a fluid under the influence of a mechanical force (forced convection) or of gravity force (free convection). At solid surfaces the velocity of fluids is zero and as a consequence only ***diffusion*** contributes to the flux at the electrode-electrolyte interface. This allows us to write the following expression for the flux of a minor ionic species B:

$$N_{\mathrm{B}} = -D_{\mathrm{B}} \left.\frac{\mathrm{d}c_{\mathrm{B}}}{\mathrm{d}y}\right|_{y=0} \qquad (\mathrm{mol/m^2s}) \tag{4.75}$$

© 2007, first edition, EPFL Press

Table 4.19 Diffusion coefficients of neutral and ionic species in water, T = 25 °C (adapted from ref. [5])

Anion	$10^9\ D$ ($m^2\ s^{-1}$)	Cation	$10^9\ D$ ($m^2\ s^{-1}$)	Molecule	$10^9\ D$ ($m^2\ s^{-1}$)
OH^-	5.24	H^+	9.32	H_2	4.06
Cl^-	2.03	Na^+	1.33	O_2	2.51
HS^-	1.73	Ag^+	1.65	N_2	2.34
NO_2^-	1.91	Ca^{2+}	0.79	H_2O_2	1.20
NO_3^-	1.90	Fe^{2+}	0.72	CO_2	1.95
HCO_3^-	1.18	Ni^{2+}	0.67		
HSO_4^-	1.33	Cu^{2+}	0.72		
SO_4^{2-}	1.07	Zn^{2+}	0.72		
CrO_4^{2-}	1.12	Pb^{2+}	0.95		
MoO_4^{2-}	0.99	Fe^{3+}	0.61		
$Fe(CN)_6^{3-}$	0.90	Cr^{3+}	0.59		
$Fe(CN)_6^{4-}$	0.73	Al^{3+}	0.54		

In this equation, *Fick's first law*, N_B represents the flux perpendicular to the electrode surface and D_B is the diffusion coefficient. Table 4.19 gives values of diffusion coefficient for a number of neutral and ionic species in dilute aqueous solution.

Figure 4.20 shows the concentration profile of a reactant (a) and of a product (b) participating in an electrode reaction. The gradient is positive in the first case, and negative in the second. By convention, cathodic and anodic current densities are, respectively, negative and positive. Considered in terms of Faraday's law, we thus obtain:

$$i_a = -n\ F\ N_B \qquad \text{(reactants)} \tag{4.76}$$

$$i_a = +n\ F\ N_B \qquad \text{(products)} \tag{4.77}$$

$$i_c = +n\ F\ N_B \qquad \text{(reactants)} \tag{4.78}$$

$$i_c = -n\ F\ N_B \qquad \text{(products)} \tag{4.79}$$

The integration of (4.75) over the interval $y = 0$ to $y = \delta$ yields

$$N_B = -D_B \frac{c_{B,b} - c_{B,s}}{\delta} \tag{4.80}$$

where $c_{B,b}$ represents the concentration of B in the bulk solution, $c_{B,s}$ its concentration at the electrode surface and δ designates the thickness of the Nernst diffusion layer.

© 2007, first edition, EPFL Press

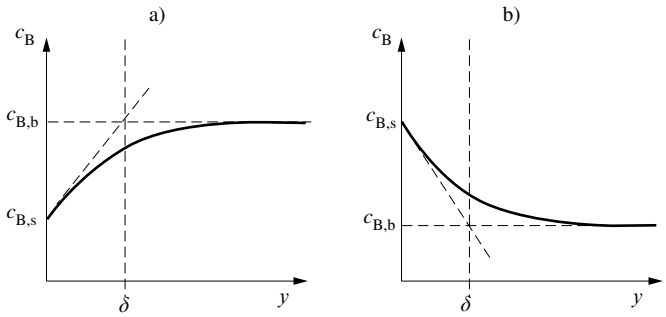


Figure 4.20 Concentration profiles in the electrolyte near the electrode surface: (a) reactant, (b) product.

Nernst diffusion layer

Figure 4.20 defines the Nernst diffusion layer, whose limit corresponds to the intersection of two straight lines, the first representing the constant concentration found in the bulk electrolyte, and the second corresponding to the tangent of the concentration profile at $y = 0$. The Nernst diffusion layer therefore does not refer to the true diffusion layer, which is larger. It represents a fictitious quantity. Its usefulness lies in the fact that it permits to greatly simplify the modeling of transport phenomena at electrodes.

In reality, as one moves away from the interface towards the bulk solution, the contribution of convection to transport increases while that of diffusion decreases. Rather than treating simultaneously transport by diffusion and convection, the Nernst model makes a clear separation between the two transport mechanisms: a total absence of convection inside the Nernst diffusion layer ($y < \delta$), and an absence of diffusion outside the Nernst diffusion layer ($y > \delta$). The intensity of convection affects the flux at the electrode by fixing the thickness of the Nernst diffusion layer. For the remainder of this book, the Nernst diffusion layer will simply be called the ***diffusion layer***.

Limiting current density

The limiting current density corresponds to the maximum reaction rate permitted by mass transport. According to (4.80), an electrode reaction reaches its limiting rate when the concentration of a *reactant B* drops to zero at the interface (4.81) or when the concentration of a *product B*, $c_{B,s}$, reaches its point of saturation, $c_{B,sat}$, at the interface (4.82).

$$i_l = \pm nF D_B \frac{c_{B,b}}{\delta} \tag{4.81}$$

$$i_l = \pm nF D_B \frac{c_{B,sat} - c_{B,b}}{\delta} \tag{4.82}$$

© 2007, first edition, EPFL Press

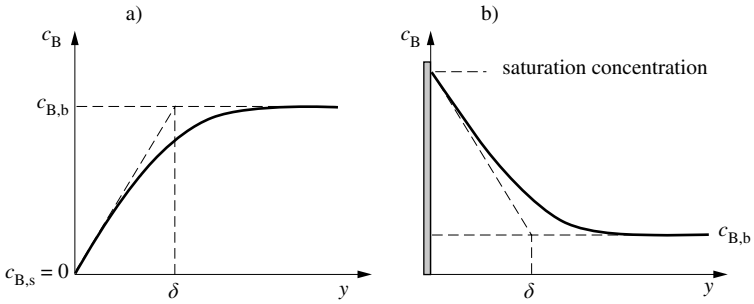


Figure 4.21 Concentration profiles near electrode surface at the limiting current : (a) reactant, (b) product.

In corrosion, the equation (4.81) normally describes the maximum rate of a reaction limited by the transport of an oxidizing agent, such as dissolved molecular oxygen, while the expression (4.82) is applicable to partial anodic reactions limited by the transport of corrosion products. Figure 4.21 shows these two situations.

Concentration overpotential

Consider the case of cupric ions present as a minor ionic species in a solution of sulfuric acid in contact with metallic copper. In the absence of dissolved oxygen, the open circuit potential of the copper electrode is equal to its reversible potential. By replacing the activity of the Cu^{2+} ions by the concentration, this gives

$$E_{(i=0)} = E_{rev} = E^0 + \frac{RT}{2F} \ln c_{Cu2+,b} \tag{4.83}$$

At equilibrium, the Cu^{2+} ion concentration is uniform: $c_{Cu2+,s} = c_{Cu2+,b}$. However, when a cathodic potential is applied copper starts to deposit and a concentration gradient forms near the electrode surface, because the reaction consumes ions. The surface concentration of cupric ion $c_{Cu2+,s}$ decreases. On the other hand, if the quantity of solution is sufficiently large, $c_{Cu2+,b}$ stays constant. By assuming that the rate of charge transfer is very rapid, the Nernst equation can still be applied at the interface, but we must now replace $c_{Cu2+,b}$ by $c_{Cu2+,s}$.

$$E = E^0 + \frac{RT}{2F} \ln c_{Cu2+,s} \tag{4.84}$$

The value of $c_{Cu2+,s}$ will vary as a function of current density and convection conditions. By introducing the overpotential $\eta = E - E_{rev}$, we obtain, with (4.83) and (4.84):

$$\eta = \frac{RT}{2F} \ln \frac{c_{Cu2+,s}}{c_{Cu2+,b}} \tag{4.85}$$

© 2007, first edition, EPFL Press

The overpotential defined by (4.85) is exclusively due to the local variation of concentration at the electrode surface, thus its name ***concentration overpotential***. In certain works (particularly older ones), the term *diffusion overpotential* is also found—a term that underscores the importance of diffusion at the electrode surface.

With (4.80) and (4.78) and dividing by (4.81) we obtain for B = Cu^{2+}:

$$\frac{c_{\mathrm{Cu2+,s}}}{c_{\mathrm{Cu2+,b}}} = 1 - \frac{i}{i_1} \tag{4.86}$$

Introducing this expression into (4.85) yields a relation between the current density and the concentration overpotential.

$$\eta = \frac{RT}{2F} \ln\left(1 - \frac{i}{i_1}\right) \tag{4.87}$$

$$i = i_1\left(1 - \exp\left(\frac{2F}{RT}\eta\right)\right) \tag{4.88}$$

For an arbitrary deposition reaction, involving n electrons, the equation becomes:

$$i = i_1\left(1 - \exp\left(\frac{nF}{RT}\eta\right)\right) \tag{4.89}$$

Figure 4.22 shows in schematic form the polarization curve corresponding to equation (4.89). The flat horizontal section indicates a limiting current, because the reaction rate no longer depends on the potential but only on the rate of mass transport.

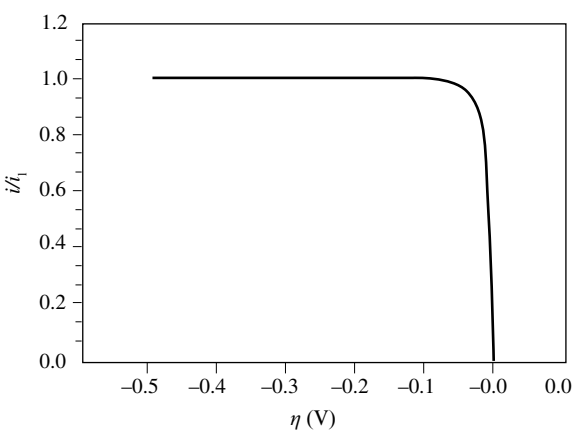


Figure 4.22 Theoretical polarization curve for cathodic deposition with only concentration overvoltage.

© 2007, first edition, EPFL Press

When the concentration of the Cu^{2+} ions at the cathode surface reaches zero, the concentration overpotential, according to (4.85), becomes infinite: $c_{Cu2+,s} = 0$, $\eta \to \infty$. In practice, such a situation is never reached because beyond a certain potential, additional electrode reactions will become possible, such as for example the reduction of protons leading to hydrogen formation. In the polarization curve the onset of additional reactions limits the extent of the limiting current plateau.

Reactions under mixed control

The rate–limiting step of electrode reactions often changes depending on current density and applied potential. For example, the electrodeposition of copper is limited by charge transfer at low overvoltage and by mass transport at high overvoltage. In this subsection we present a way to take into account both effects. In order to simplify the calculations, we assume that the anodic partial reaction can be neglected. Equation (4.35) then becomes:

$$i_c = -i_0 \frac{c_{Cu2+,s}}{c_{Cu2+,b}} \exp\left(-\frac{\eta}{\beta_c}\right) \tag{4.90}$$

With (4.86) we obtain

$$i_c = -i_0 \left(1 - \frac{i_c}{i_l}\right) \exp\left(-\frac{\eta}{\beta_c}\right) \tag{4.91}$$

and after rearrangement:

$$i_c = -\frac{i_0 \exp\left(-\frac{\eta}{\beta_c}\right)}{1 - \frac{i_0}{i_l} \exp\left(-\frac{\eta}{\beta_c}\right)} \tag{4.92}$$

By dividing this equation by the limiting current density, i_l, we obtain its dimensionless form:

$$\frac{i_c}{i_l} = -\frac{\frac{i_0}{i_l} \exp\left(-\frac{\eta}{\beta_c}\right)}{1 - \frac{i_0}{i_l} \exp\left(-\frac{\eta}{\beta_c}\right)} \tag{4.93}$$

For weak cathodic overpotential, and if $i_0 << |i_l|$, the second term of the denominator is negligible and equation (4.93) simply represents the cathodic term of the Butler–Volmer equation (4.36). For high cathodic overpotentials ($\eta/\beta_c << -1$), the exponential terms dominate and, as a result, the ratio i_c/i_l tends toward unity. In this case, mass transport controls the reaction rate.

© 2007, first edition, EPFL Press

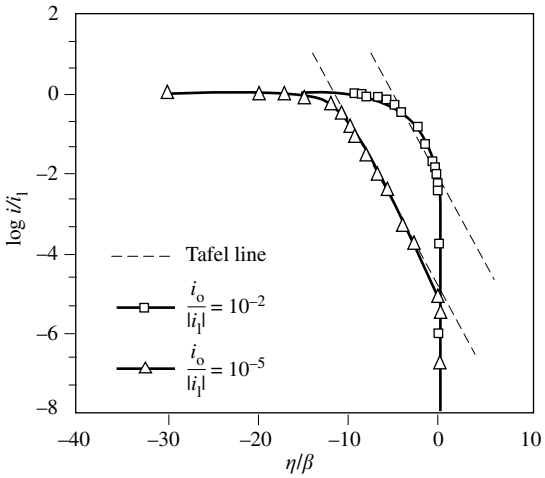


Figure 4.23 Theoretical polarization curves for cathodic depositon under mixed control (log scale); $\alpha = 0.5$, $n = 2$.

Figure 4.23 shows, on a logarithmic scale, the variation of the dimensionless cathodic current density as a function of the dimensionless overpotential η/β_c for two values of the ratio $i_0/|i_l|$. The region under charge–transfer control spreads as the ratio $i_0/|i_l|$ becomes smaller. Presentation of the same data on a linear scale (Figure 4.24) shows that a decrease in the ratio $i_0/|i_l|$ shifts the curve towards higher overpotentials and shortens the plateau of the limiting current. On the other hand, if the ratio $i_0/|i_l|$ becomes large, the behavior of the system approaches that described by (4.88) for pure concentration overvoltage.

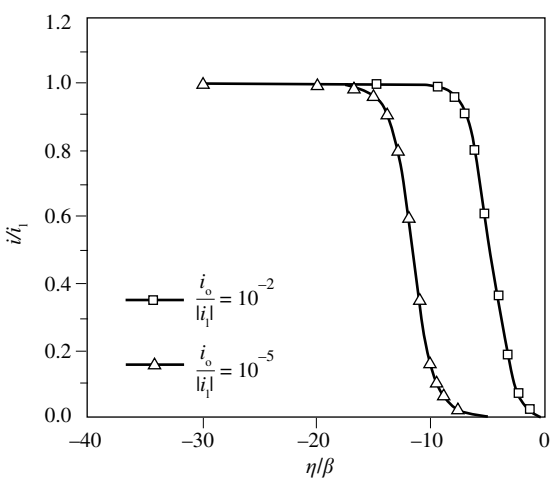


Figure 4.24 Theoretical polarization curves for cathodic depositon under mixed control (linear scale); $\alpha = 0.5$, $n = 2$.

© 2007, first edition, EPFL Press

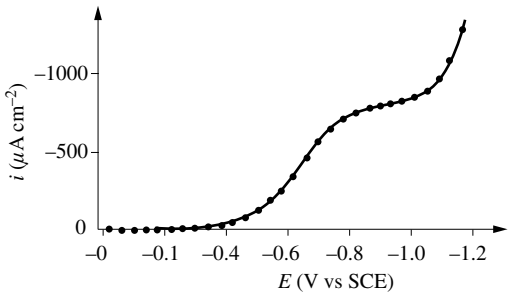


Figure 4.25 Cathodic polarization curve for oxygen reduction on platinum in 3% NaCl on rotating disk electrode at 1000 rpm.

Equations (4.91) to (4.93) can be applied to any cathodic partial reaction for which the charge-transfer step obeys the Butler-Volmer equation. In corrosion, *oxygen reduction* is often under mixed control. Figure 4.25 shows the cathodic polarization curve for oxygen reduction, measured on a platinum electrode [6]. The shape of the curve suggests a relatively low value for the ratio $i_0/|i_l|$.

Generalization

The theory can be generalized by also taking into account the anodic term of the Butler-Volmer equation. If the anodic partial reaction is controlled by charge transfer, and if the cathodic reaction is under mixed control, the polarization curve is described by equation (4.94).

$$i = i_a + i_c = i_0 \exp\left(\frac{\eta}{\beta_a}\right) - i_0\left(1 - \frac{i_c}{i_l}\right)\exp\left(-\frac{\eta}{\beta_c}\right) \tag{4.94}$$

By setting:

$$\frac{i}{i_l} = \frac{i_a}{i_l} + \frac{i_c}{i_l} \tag{4.95}$$

we obtain, with (4.93):

$$\frac{i}{i_l} = \frac{i_0}{i_l}\exp\left(-\frac{\eta}{\beta_a}\right) - \frac{i_0}{i_l}\frac{\exp\left(-\frac{\eta}{\beta_c}\right)}{1 - \frac{i_0}{i_l}\exp\left(-\frac{\eta}{\beta_c}\right)} \tag{4.96}$$

4.3.2 Convective transport

The thickness of the diffusion layer and therefore the limiting current density vary as a function of convection conditions. There are two forms of convection:

© 2007, first edition, EPFL Press

- forced convection, due to the action of an external force, for example a pump;
- free convection, due to the action of the buoyancy force on a fluid whose density is not homogeneous.

Forced convection

Consider a fluid moving through a pipe in the laminar flow regime. The wall of the pipe contains an electrode, located at a certain distance from the entry (Figure 4.26). The flow rate at the wall is zero. In the vicinity of the walls, viscous forces slow down the fluid as soon as it enters the pipe. Thus a gradient in flow rate is established across a layer referred to as *the hydrodynamic boundary layer*. Its thickness increases with the distance from the inlet. The boundary layers of opposing walls eventually meet after a distance L_h, called the *hydrodynamic entrance length*. From this point onward, the flow profile is observed to be parabolic. For a tube, L_h has a value of about 70 times its diameter.

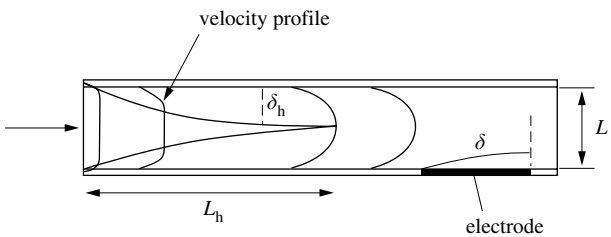


Figure 4.26 Laminar flow of the electrolyte in a pipe: δ_h = hydrodynamic boundary layer, δ = diffusion layer.

If an electrode reaction, controlled by mass transport, takes place on an electrode embedded in the wall, a diffusion layer is established. Its thickness increases as one moves downstream, because the concentration change due to the electrochemical reaction affects an ever-increasing volume. The current density on the electrode is therefore not uniform but decreases towards the downstream end of the electrode. Thus, the longer the electrode, the more the average limiting current density becomes smaller.

This example highlights the need for distinguishing between the hydrodynamic boundary layer, which is given by the flow velocity profile, and the diffusion layer, which depends on the concentration profile at the electrode. To calculate the rate of mass transport under laminar flow conditions, one must first evaluate the velocity profile near the electrode. How this can be accomplished is described in detail elsewhere [7,8].

Dimensionless correlations

The rate of mass transport at electrodes, in laminar or turbulent regimes, has been studied both theoretically and experimentally for many geometries. In order to reduce the number of variables, the results are normally reported in dimensionless form.

© 2007, first edition, EPFL Press

In forced convection systems, the ***Sherwood number***, *Sh*, characterizes the dimensionless rate of mass transport at the electrodes.

$$Sh = \frac{|i_1| L}{n F D c_b} \tag{4.97}$$

Here, i_l is the limiting current density, L is a characteristic length of the system (such as the diameter of a pipe), c_b is the bulk concentration of the rate limiting species in the solution, and D is its diffusion coefficient. Equation (4.81) allows us to write:

$$Sh = \frac{L}{\delta} \tag{4.98}$$

This definition provides a simple relation between the Sherwood number and the thickness of the diffusion layer, δ.

In a forced-convection system, the Sherwood number is a function of the Reynolds number, *Re*, and the Schmidt number, *Sc*.

$$Sh = f(Re, Sc) \tag{4.99}$$

The ***Reynolds number*** may be viewed as a dimensionless flow rate:

$$Re = \frac{u L}{\nu} \tag{4.100}$$

where u is the linear flow rate of the fluid (m/s), L is a characteristic length of the system, and ν is the kinematic viscosity of the fluid ($m^2\ s^{-1}$), which is equal to the viscosity divided by the density.

For rotating electrodes, the linear velocity is equal to $u = \omega r$ where ω refers to the angular velocity (rad/s) and r is the radius. The Reynolds number, therefore, becomes:

$$Re = \frac{\omega r L}{\nu} \tag{4.101}$$

The ***Schmidt number*** expresses to the ratio between the kinematic viscosity ν and the diffusion coefficient:

$$Sc = \frac{\nu}{D} \tag{4.102}$$

Typical values for aqueous solutions are: $\nu \approx 10^{-6}\ m^2\ s^{-1}$, $D \approx 10^{-9}\ m^2\ s^{-1}$. This yields a Schmidt number of $Sc \approx 1000$. Sometimes the Schmidt number and the Reynolds number are combined in a new variable, called the ***Péclet number***:

$$Pe = Re\ Sc \tag{4.103}$$

Table 4.27 gives dimensionless mass transport correlations for different geometries and flow regimes. More complete compilations can be found in the literature [9,10].

© 2007, first edition, EPFL Press

Table 4.27 Transport correlations for forced convection systems

Geometry	Flow	Characteristic length	*Sh* (average value, $Sc \geq 1000$)
Pipe, smooth walls	turbulent $Re>3000$	diameter, D_h	$0.0115\ Re^{7/8}Sc^{1/3}$
Pipe, smooth walls	laminar, $Re< 2000$ fully developed velocity profile $D_h/L_x > 1.85$	diameter, D_h	$1.85\ Re^{1/3}Sc^{1/3}(D_h/L_x)^{1/3}$ (L_x = length of electrode)
Pipe, smooth walls	laminar, $Re< 2000$ fully developed velocity profile, $D_h/L_x \rightarrow 0$	diameter, D_h	3.66
Rotating disk	laminar $Re < 2.7\times10^5$	radius	$0.62\ Re^{1/2}Sc^{1/3}$
Rotating disk	turbulent $8.9\times10^5 < Re < 1.18\times10^7$	radius	$0.0117\ Re^{0.896}\ Sc^{0.249}$
Rotating hemisphere	laminar $Re < 1.5\times10^4$	radius	$0.474\ Re^{1/2}Sc^{1/3}$
Rotating cylinder	turbulent $10^3 < Re < 10^5$	diameter	$0.079\ Re^{0.7}Sc^{-0.35}$

Rotating disk electrode

Rotating disk electrodes (Figure 4.28) are extensively used in electrochemistry and corrosion for the study and control of mass-transport phenomena, because it allows the operator to easily and reproducibly vary the transport conditions. Furthermore, on a rotating disk electrode the mass transport rate is uniform over the entire electrode surface, meaning that the diffusion layer thickness does not depend on radius. A flat disk, enclosed by insulating material, rotates in an electrolyte under laminar flow conditions (typically 50 to 5000 rotations per minute). The rotation draws fluid towards the disk and projects it outward towards the edge by the action of tangential forces, thereby creating a spiral-shaped motion. Figure 4.29 schematically shows an airtight experimental set-up consisting of a rotating-disk electrode, a reference electrode with a Luggin capillary, and a counter electrode separated from the main compartment by a frit. The details of this type of set-up may vary according to the problem under study: A frit is not always needed to separate the counter electrode and not all experiments need to be carried out in an airtight, de-aerated cell. Certain researchers position the Luggin capillary away from the disk in order to avoid any perturbation to the hydrodynamic conditions. In that case, it is advisable to

© 2007, first edition, EPFL Press

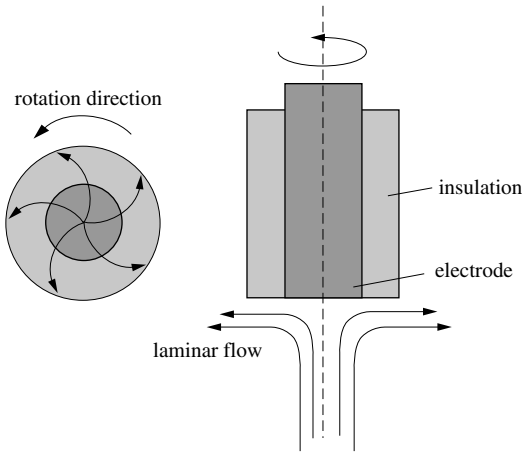


Figure 4.28 Laminar flow of the electrolyte near the surface of a rotating disk electrode.

perform a mathematical correction for the ohmic drop (Section 11.5). For a virtual reference electrode, situated at infinite distance, we can calculate the ohmic drop with equation (4.104), where r refers to the radius of the disk [11]:

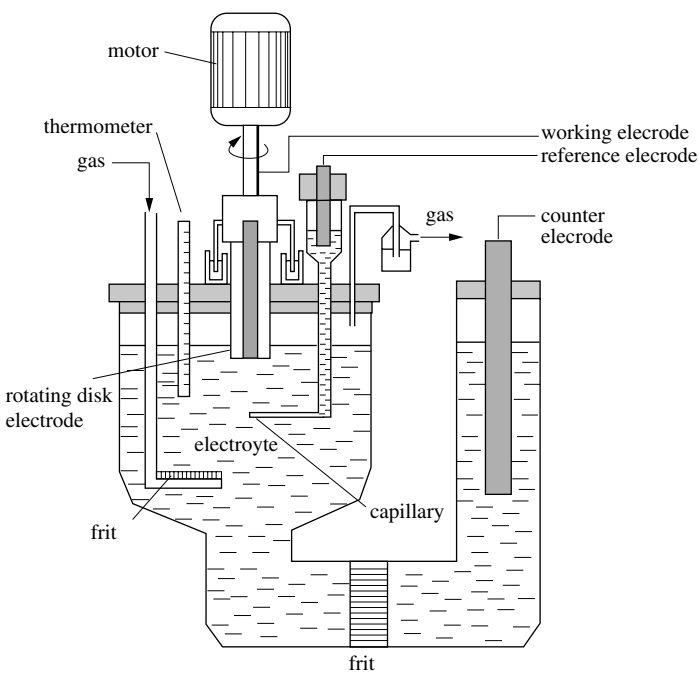


Figure 4.29 Electrochemical cell with a rotating disk electrode.

© 2007, first edition, EPFL Press

$$\Delta\Phi_{\Omega} = \frac{I}{4\pi r \kappa} \tag{4.104}$$

The theory describing the mass transport to a rotating-disk electrode is due to Levich [8], who first presented an analytical solution for the flux to the electrode under laminar flow conditions ($Re < 2.7\text{x}10^5$). For the reaction of a minor ionic species B, the limiting current is given by equation (4.105), know as the ***Levich equation***:

$$|i_l| = 0.62\, n\, F\, c_{B,b}\, D_B^{2/3}\, \nu^{-1/6}\, \omega^{1/2} \tag{4.105}$$

Here $\omega = 2\pi f$ represents the angular velocity in rad s^{-1} (f is the frequency in s^{-1}), ν is the kinematics viscosity in $cm^2\ s^{-1}$, D_A is the diffusion coefficient in $cm^2\ s^{-1}$, $c_{B,b}$ is the concentration of B in the bulk in mole cm^{-3}, and i_l is the limiting current density in A cm^{-2}.

By combining (4.105) and (4.81), we obtain an expression for the thickness of Nernst diffusion layer.

$$\delta = 1.61\, D_B^{1/3}\, \nu^{1/6}\, \omega^{-1/2} \tag{4.106}$$

This equation confirms that the diffusion layer thickness is independent of the disk radius.

According to the Levich equation (4.105) the limiting current density for a rotating disk electrode is proportional to $c_B\, D_B^{2/3}$. On the other hand, the Nernst model (equation 4.81) indicates that the limiting current is proportional to the product $c_B\, D_B$. By combining these two equations we find that δ has a $D_B^{1/3}$ dependence. This result reveals the artificial character of the Nernst diffusion layer model. Every dissolved species that reacts in an electrochemical system has its own specific Nernst diffusion layer.

Electrode reactions under mixed control

The rotating-disk electrode is well suited for the study of the kinetics of electrochemical reactions under mixed control, because it permits one to separate the contribution of mass transport and of interface kinetics to the overall reaction rate.

When a dissolved species cathodically reacts at potentials far from equilibrium ($\eta/\beta_c << 1$) the measured current density i is equal to the cathodic partial current density, and the polarization curve will be described by equation (4.91). For infinitely fast transport rates ($i_l \to \infty$) , the cathodic partial current density corresponds to i_t:

$$i_t = -i_0 \exp\left(-\frac{\eta}{\beta_c}\right) \tag{4.107}$$

Combining (4.91) and (4.107) gives after some rearrangement:

$$\frac{1}{i} = \frac{1}{i_t} + \frac{1}{i_l} \tag{4.108}$$

© 2007, first edition, EPFL Press

According to the Levich equation (4.105), the limiting current density of a rotating-disk electrode can be written as:

$$|i_l| = \frac{\omega^{1/2}}{K_\omega} \tag{4.109}$$

where

$$\frac{1}{K_\omega} = 0.62\, n\, F\, c_{B,b}\, D_B^{2/3}\, \nu^{-1/6}$$

With (4.108), we thus find:

$$\frac{1}{|i|} = \frac{1}{|i_t|} + \frac{K_\omega}{\omega^{1/2}} \tag{4.110}$$

If $1/i$ is plotted as a function of $1/\omega^{1/2}$, the intersection of the straight line with the y-axis gives $1/i_t$ and from the slope one can get K_ω (Figure 4.30). Thus one obtains information on both, the rate of mass transport and the rate of the interfacial reaction.

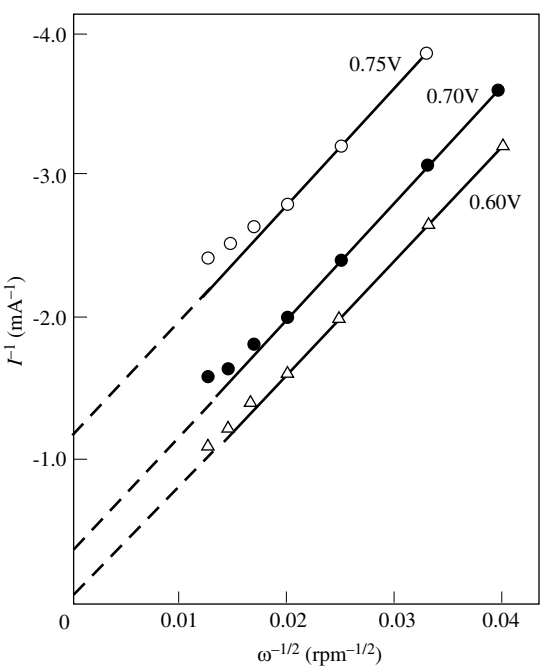


Figure 4.30 Kouteki-Levich plot for reduction of $Fe(CN)_6^{3-}$ to $Fe(CN)_6^{4-}$ on a 0.385 cm^2 platinum rotating disk electrode under mixed control in 0.5M K_2SO_4, at 28 °C. The reciprocal of the measured current is shown as a function of $\omega^{-1/2}$ for three potentials [12].

© 2007, first edition, EPFL Press

Free convection

In electrochemical systems, consumption or accumulation of reactants and products at the electrode change the local density of the electrolyte in the diffusion layer. This causes buoyancy forces to act on the solution near the electrode and leads to the establishment of free convection flow conditions. More rarely, temperature gradients may also be at the origin of local density differences the electrolyte.

The ions resulting from anodic metal dissolution are present at higher concentration near the electrode surface than in the bulk, and this leads to a locally higher density of the electrolyte. In an unstirred cell with vertical electrodes, the resulting buoyancy force (Figure 4.31) causes the electrolyte in the anodic diffusion layer to move downwards. In an analogous way, an electrodeposition reaction results in an upward movement of the electrolyte because it lowers the metal ion concentration in the vicinity of the electrode.

A theoretical analysis of the deposition of a metal B on a vertical electrode of length L, in the laminar flow regime, yields the following expression for the average limiting current density [10]:

$$|i_1| = 0.67\, n\, F\, g^{1/4}\, \alpha'^{1/4}\, D_B^{3/4}\, \nu^{-1/4}\, c_{B,b}^{5/4}\, L^{3/4} \tag{4.111}$$

where g represents the gravitational constant (g = 9.80665 m s^{-2}) and α' is a proportionality factor relating the electrolyte density to the concentration of B. Interestingly in this example, the limiting current does not vary linearly with the concentration, but rather according to a 5/4 power expression.

By analogy with forced convection, we present the rate of mass transport caused by free convection in a dimensionless form. For a vertical electrode immersed in a fluid subject to free convection in the laminar flow regime ($Sc \cdot Gr < 10^{12}$), we find for the average Sherwood number:

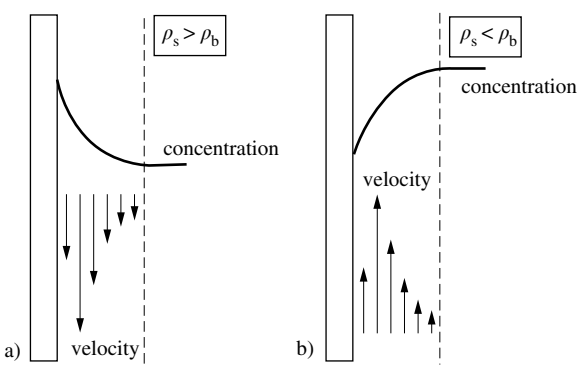


Figure 4.31 Velocity profiles near an electrode surface under conditions of free convection. The density of the electrolyte at the surface, ρ_s, can be either greater (a) or smaller (b) than that of the bulk, ρ_b.

© 2007, first edition, EPFL Press

$$Sh = 0.67\ Sc^{1/3}\, Gr^{1/4} \tag{4.112}$$

where Gr represents the ***Grashof number***,

$$Gr = \frac{g\ \Delta\rho\ L^3}{\rho_b\, \nu^2} \tag{4.113}$$

In this equation, the term $\Delta\rho = |\rho_s - \rho_b|$ corresponds to the difference, in absolute value, of the volumetric mass of the electrolyte between the interface (ρ_s) and the bulk (ρ_b), L is a characteristic length (in this case the height of the electrode) and ν represents the kinematic viscosity. For a binary electrolyte of concentration c_b, we find:

$$\frac{\Delta\rho}{\rho_b} = \alpha'\, \Delta c_i \tag{4.114}$$

where α' is a proportionality factor. For a multi-component electrolyte:

$$\frac{\Delta\rho}{\rho_b} = \sum \alpha'_i\, \Delta c_i \tag{4.115}$$

$$\Delta c_i = \left| c_{i,s} - c_{i,b} \right|$$

where α'_i represents a proportionality factor between the concentration of a species i and the density of the solution. Evaluation of the ratio $\Delta\rho/\rho_b$ by (4.115) can be problematic, because it requires knowledge of the concentration profiles for all the species present [10].

Horizontal electrodes

For horizontal electrodes, it is necessary to make a distinction between two basic situations, according to the direction of the density gradient in the vicinity of the electrode (Figure 4.32).

If the density of the fluid decreases with height, gravity forces do not lead to free convection. This situation typically is found for metal deposition on a horizontal electrode, positioned face down, or for dissolution of an electrode positioned face up. In these cases, the transport of reactants and products is due to non steady state diffusion, and the current therefore never reaches a constant value. Equation (4.116) indicates the variation of the limiting current density as a function of the reaction time t, for a reaction carried out at constant potential and controlled by the diffusion of a species B towards the electrode.

$$\left| i_l \right| = n\, F\, c_{B,b}\, \frac{D_B^{1/2}}{\pi^{1/2} t^{1/2}} \tag{4.116}$$

Equation (4.116) will be discussed in more detail in Chapter 5.

© 2007, first edition, EPFL Press

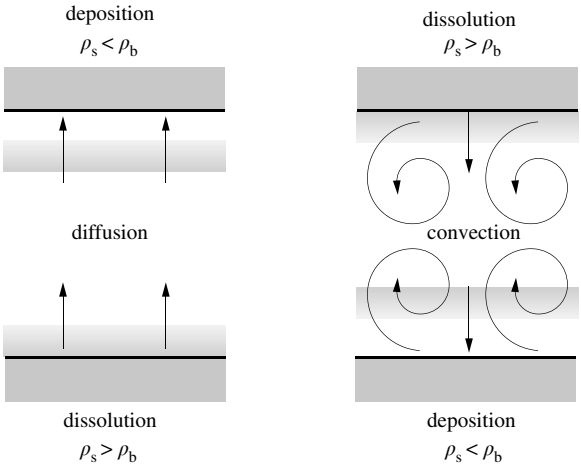


Figure 4.32 Free convection at the surface of a horizontal electrode positioned either upwards or downwards. The density of the electrolyte at the surface, ρ_s, can be either greater or smaller than that of the bulk, ρ_b.

If the fluid density near a horizontal electrode increases with height, eddies will form in the fluid under the influence of gravity. This situation is encountered during the dissolution of an electrode in a face-down position or for the deposition of a metal on an electrode positioned face up. Empirical dimensionless correlations of the type $Sh = f(Gr,Sc)$ can be used to describe the transport of the reactant and products under these conditions (Table 4.33). From a practical point of view, it is important to realize that it usually takes a few minutes to establish stable free convection conditions, because the electrochemical reaction and fluid flow are mutually interdependent.

Table 4.33 Dimensionless mass-transport correlations for free convection systems [9].

Electrode	Flow	Characteristic length	Sh
Vertical plane electrode	laminar	height	$0.67\ (Sc\ Gr)^{1/4}$
Vertical plane electrode	turbulent	height	$0.31\ (Sc\ Gr)^{0.28}$
Horizontal plane electrode facing upwards	laminar	surface/perimeter	$0.54\ (Sc\ Gr)^{1/4}$
Horizontal plane electrode facing downwards	turbulent	surface/perimeter	$0.15\ (Sc\ Gr)^{1/3}$

Convection induced by gas bubbles

An electrode that undergoes corrosion in an acidic environment often gives off hydrogen bubbles. These bubbles stir the solution in the immediate vicinity of the

© 2007, first edition, EPFL Press

electrode, thus modifying the mass transport conditions. Equation (4.117) expresses the limiting current density as a function of the rate of hydrogen release, v_g (cm^3 cm^{-2} s), for the reaction of a species B, on a vertical electrode [10].

$$|i_l| = 0.03\, n\, F\, c_{B,b}\, v_g^{0.53} \tag{4.117}$$

In this equation, valid for the range $3 \times 10^{-4} < v_g < 2 \times 10^{-2}$, $c_{B,b}$ represents the bulk concentration of the species B in mol cm^{-3}, and i_l corresponds to the limiting current density, in A cm^{-2}.

4.3.3 Transport by migration

So far we considered mass transport by convection and diffusion, neglecting a possible influence of potential gradients present in the electrolyte. The equations presented in the preceding sections therefore are strictly applicable only to neutral species, or to minor ionic species present in a supporting electrolyte. The effect of potential gradients on the rate of transport of ionic species is particularly pronounced in solutions where concentration gradients are absent or when two types of ions are present in comparable quantities. We therefore need to have a closer look at the transport mechanisms in electrolytes in presence of a potential gradient.

Transport equations

Three phenomena contribute to the transport of ions in an electrolyte:

- diffusion, due to a concentration gradient;
- migration, due to a gradient in the electric potential;
- convection, due to fluid motion.

The ***flux equation*** thus has the following form:

$$N_B = -D_B \nabla c_B - z_B\, F\, u_B c_B \nabla \Phi + c_B v \tag{4.118}$$

where N_B represents the flux of ions B (mol m^{-2} s^{-1}), ∇c_B is the concentration gradient of B (mol cm^{-4}), $\nabla \Phi$ is the potential gradient (V m^{-1}), z_B is the charge on B, u_B is the mobility of B (m^2 mol J^{-1} s^{-1}), and v is the flow velocity of the solution (m s^{-1}).

Table 4.34 Transport equations.

Type	Equation
Flux	$N_B = -D_B \nabla c_B - z_B F\, u_B c_B \nabla \Phi + c_B v$
Current density	$i = F\, \Sigma z_i\, N_i$
Electroneutrality	$\Sigma z_i c_i = 0$
Mass conservation	$\frac{\partial c_B}{\partial t} = -\nabla \bullet N_B + R_B$

© 2007, first edition, EPFL Press

The movement of ionic charge in the electrolyte corresponds to an electrical current. The magnitude of the corresponding current density is given by the sum of all the ionic fluxes, multiplied by the corresponding ionic charge $z_i F$.

$$i = F \Sigma z_i N_i \tag{4.119}$$

Electrolyte solutions are electrically neutral, with the exception of the electric double layer region. However, because the thickness of the double layer is much smaller than that of the diffusion layer, its effect can usually be neglected when discussing transport phenomena. The equation (4.120) expresses the neutrality condition for the electrolyte.

$$\Sigma z_i c_i = 0 \tag{4.120}$$

A mass balance of a species B over a given volume element in the solution results in the relation (4.121), where $\nabla \bullet N_B$ (mol m^{-3} s^{-1}) represents the divergence of the flux of B, and R_B (mol m^{-3}s^{-1}) indicates the rate of formation (or loss) of B by a homogeneous chemical reaction.

$$\frac{\partial c_B}{\partial t} = -\nabla \bullet N_B + R_B \tag{4.121}$$

Mobility

The mobility u_B characterizes the velocity (m s^{-1}) that one mole of ions attains under the applied force of 1 Newton. It is expressed in units of mole m s^{-1} N^{-1} or mole m^2 J^{-1} s^{-1}, and its definition remains valid regardless of the nature of the applied force. Often, the force is due to an electric field. In some texts, this has led to the definition of the ***electrical mobility*** u_e as the velocity that an ion attains in a potential gradient of 1 V m^{-1}. The units for u_e are m^2 V^{-1} s^{-1}. Mobility and electrical mobility are related by the expression:

$$u_{eB} = z_B F u_B \tag{4.122}$$

In a general way, the flux of the ions B in the absence of convection is proportional to the gradient of the electrochemical potential $\nabla \mu_B^*$.

$$N_B = -c_B u_B \nabla \mu_B^* = -c_B u_B \nabla(\mu_B + z_B F \Phi) \tag{4.123}$$

We set for the electrochemical potential $\mu_B^* = \mu_B + zF\phi$ (Chapter 2) and for the chemical potential:

$$\mu_B = \mu_B^0 + RT \ln f_B c_B \tag{4.124}$$

By assuming the term $\nabla \ln f_B \approx 0$ to be negligible, equation (4.123) thus becomes:

$$N_B = -c_B u_B RT \nabla \ln c_B - c_B u_B z_B F \nabla \Phi \tag{4.125}$$

© 2007, first edition, EPFL Press

With $\nabla \ln c_B = (\nabla c_B / c_B)$ we finally get:

$$N_B = -u_B RT \nabla c_B - c_B u_B z_B F \nabla \Phi \tag{4.126}$$

Comparison of (4.126) with the first two terms of the flux equation (4.118) allows us to establish the ***Nernst-Einstein equation*** (4.127), which relates the mobility of a species B, u_B, to its diffusion coefficient, D_B.

$$D_B = u_B RT \tag{4.127}$$

The above derivation indicates that the flux equation (4.118) is strictly applicable only to dilute solutions for which the gradients of activity coefficients are negligible. The theory, based on the four equations summarized in Table 4.34, is therefore often referred to as *dilute solution transport theory*. Dilute solution transport theory is commonly used in practice to describe transport phenomena in electrolytes.

Conductivity

Ohm's law defines the conductivity κ ($\Omega^{-1}m^{-1}$) of an electrolyte in the absence of concentration gradients:

$$i = -\kappa \nabla \Phi \tag{4.128}$$

Some authors define the *molar conductivity* Λ_{AB} of the salt *AB* by the ratio $\Lambda_{AB} = \kappa / c_{AB}$, or the *equivalent conductivity* Λ'_B of an ion as $\Lambda'_B = z_B F u_{eB}$. Because they easily lead to confusion, these quantities will not be used in this text.

According to (4.118) and if there is no concentration gradient ($\nabla c_i = 0$), the current density in the electrolyte is equal to:

$$i = -F^2 \Sigma z_i^2 u_i c_i \nabla \Phi + F \Sigma z_i c_i v \tag{4.129}$$

The electroneutrality condition (4.120) cancels out the second term. By comparison with (4.128), we find:

$$\kappa = F^2 \Sigma z_i^2 u_i c_i \tag{4.130}$$

This equation shows that the conductivity depends on the mobility of the ions present and it should increase linearly with ion concentration. As an illustration, Table 4.35 gives the conductivity of KCl solutions at different concentrations. In this case,

$$z_{K+} = -z_{Cl-} = 1; \quad c_{K+} = c_{Cl-} = c_{KCl}$$

and the equation (4.130) therefore becomes:

$$\kappa = F^2 (u_{K+} + u_{Cl-}) c_{KCl} \tag{4.131}$$

The experimental results shown in Table 4.35 obey the expression (4.131) relatively well. At high concentrations, the electrostatic interaction between ions causes a reduction in their mobility, thus resulting in a slight deviation from linearity.

© 2007, first edition, EPFL Press

Table 4.35 Conductivity of KCl solutions at 25 ° C

c_{KCl} (mol l^{-1})	κ ($\Omega^{-1}cm^{-1}$)
0.001	1.47×10^{-4}
0.01	1.41×10^{-3}
0.1	1.29×10^{-2}
1	1.11×10^{-1}

Transport number

In the absence of concentration gradients, the transport number t_B indicates the fraction of the ionic current carried by the migration of the ionic species B:

$$t_B = \frac{i_B}{i} \tag{4.132}$$

Equations (4.118) and (4.119) yield for the ionic current density resulting from the migration of B:

$$i_B = - z_B^2 F u_B^2 c_B \nabla \Phi \tag{4.133}$$

The total current density according to (4.129) is

$$i = - F^2 \Sigma z_i^2 u_i c_i \nabla \Phi \tag{4.134}$$

Combining these equations allows us to relate the transport number of B to the concentrations and mobilities of all ionic species present in the electrolyte:

$$t_B = \frac{z_B^2 u_B c_B}{\sum z_i^2 u_i c_i} \tag{4.135}$$

Two limiting cases of (4.135) are of particular interest:

- B is a minor ionic species in a supporting electrolyte;
- B is an ionic component of a binary electrolyte.

If B is a minor ionic species in presence of a supporting electrolyte, we have by definition $c_B << \sum c_i$ and as a consequence its transport number has a very small value: $t_B << 1$. This explains why migration does not significantly contribute to the rate of mass transport of minor ionic species at electrodes and to the value of the limiting current (see Sect. 4.3.1).

A *binary electrolyte* by definition contains only one type of cations and of anions. For a binary electrolyte containing a dissolved salt of formula $M_{\nu+}X_{\nu-}$ at a concentration of c_{MX}, the cation concentration is equal to $c_+ = \nu_+ c_{MX}$ and the anion concentration is $c_- = \nu_- c_{MX}$. Inserting these values into equation (4.135) gives for the transport number of the cation, t_+, and that of the anion, t_- , in a binary electrolyte:

© 2007, first edition, EPFL Press

$$t_+ = \frac{z_+ u_+}{z_+ u_+ - z_- u_-} \tag{4.136}$$

$$t_- = \frac{-z_- u_-}{z_+ u_+ - z_- u_-} \tag{4.137}$$

We conclude that in a binary electrolyte, the transport number depends only on the mobility of the ions present but not on their concentration.

Flux at the electrodes

Transport by migration contributes to the flux of reacting species at an electrode and therefore affects the limiting current density. The effect is largest in a binary electrolyte and smallest for the reaction of a minor ionic species present in a supporting electrolyte.

To estimate the effect of migration on the limiting current density we consider a specific case: the cathodic reduction of a cation M^{n+} from a binary electrolyte, such as deposition of copper from a $CuSO_4$ solution.

We assume that the Nernst diffusion layer model is applicable and therefore within the diffusion layer only diffusion and migration contribute to mass transport. Equation (4.118) then gives for the cation flux N_+ and the anion flux N_- at the electrode surface:

$$N_+ = -D_+ \frac{\mathrm{d}c_+}{\mathrm{d}y} - z_+ F u_+ c_+ \frac{\mathrm{d}\Phi}{\mathrm{d}y} = \frac{i}{nF} \tag{4.138}$$

$$N_- = -D_- \frac{\mathrm{d}c_-}{\mathrm{d}y} - z_- F u_- c_- \frac{\mathrm{d}\Phi}{\mathrm{d}y} = 0 \tag{4.139}$$

Note that the anion flux at the electrode surface is zero, because in our example only the cations undergo a charge transfer reaction. The term $\mathrm{d}\Phi/\mathrm{d}y$ in equation (4.138) can now be replaced by (4.139):

$$\frac{\mathrm{d}\Phi}{\mathrm{d}y} = -\frac{D_-}{z_- F u_- c_-} \frac{\mathrm{d}c_-}{\mathrm{d}y} \tag{4.140}$$

$$\frac{i}{nF} = -D_+ \frac{\mathrm{d}c_+}{\mathrm{d}y} + \frac{z_+ u_+ c_+ D_-}{z_- u_- c_-} \frac{\mathrm{d}c_-}{\mathrm{d}y} \tag{4.141}$$

The electroneutrality relation $c_+ z_+ = c_- z_-$ allows us to replace c_- :

$$\frac{i}{nF} = -\frac{z_- u_- D_+ - z_+ u_+ D_-}{z_- u_-} \frac{\mathrm{d}c_+}{\mathrm{d}y} \tag{4.142}$$

© 2007, first edition, EPFL Press

Application of the Nernst-Einstein equation yields:

$$\frac{i}{nF} = -\left(1 - \frac{z_+}{z_-}\right) D_+ \frac{\mathrm{d}c_+}{\mathrm{d}y} \tag{4.143}$$

Integration from $y = 0$ to $y = \delta$ gives:

$$i = -nF\left(1 - \frac{z_+}{z_-}\right) D_+ \frac{c_{+,\mathrm{b}} - c_{+,\mathrm{s}}}{\delta} \tag{4.144}$$

At the limiting current, $i = i_l$, $c_{+,\mathrm{s}} = 0$, which allows us to write:

$$i_l = -nF\left(1 - \frac{z_+}{z_-}\right) D_+ \frac{c_{+,\mathrm{b}}}{\delta} \tag{4.145}$$

A comparison of this expression with (4.81) for pure diffusion, indicates that in a binary electrolyte the influence of migration on the limiting current density shows up in the factor $(1 - z_+/z_-)$. If $z_+ = z_-$ as for $CuSO_4$, migration augments the value of the limiting current by a factor of two.

One can define the diffusion coefficient of the salt $D_\pm$ by

$$D_\pm = \frac{z_+u_+D_+ - z_-u_-D_-}{z_+u_+ - z_-u_-} \tag{4.146}$$

Introducing this expression into equation (4.142) and setting for the transport number of the anion, $t_- = 1 - t_+$, we obtain another expression for the limiting current density in a binary electrolyte:

$$i_l = -nF\frac{D_\pm}{1 - t_+}\frac{c_{+,\mathrm{b}}}{\delta} \tag{4.147}$$

This equation, equivalent to (4.144), highlights the influence of the transport number on the limiting current.

For the reduction of a cation in an electrolyte containing ions of several types, the following approximation is sometimes made:

$$i \approx -nFD_+\frac{c_{+,\mathrm{b}} - c_{+,\mathrm{s}}}{\delta} + t_+ i \tag{4.148}$$

The first term represents the flux due to diffusion, and the second term approximates the effect of migration (it would be strictly valid only in absence of concentration gradients). Rearrangement of (4.148) yields for the current density:

$$i \approx -nF\frac{D_+}{1 - t_+}\frac{c_{+,\mathrm{b}} - c_{+,\mathrm{s}}}{\delta} \tag{4.149}$$

© 2007, first edition, EPFL Press

Likewise, for the limiting current density:

$$i_l \approx -nF \frac{D_+}{1-t_+} \frac{c_{+,b}}{\delta} \tag{4.150}$$

Equation (4.150) serves to estimate the effect of migration on the limiting current density in solutions containing different types of ions. If the transport number of the reacting species is small, $t_+ << 1$, the contribution of migration is negligible.

In some cases it is useful to replace the term $D_+/(1 - t_+)$ by an ***effective diffusion coefficient*** D_{eff} that includes the contribution of migration:

$$i_l = -nF D_{eff} \frac{c_{+,b}}{\delta} \tag{4.151}$$

This approach permits for example to apply dimensionless mass transport correlations as discussed in section (4.3.2). However, the value of D_{eff} must be determined experimentally for each electrolyte composition because it depends on the concentration of all ions present. The reasoning presented here for the cathodic reduction of a cation can be easily extended to the reduction of anions and to anodic reactions.

4.3.4 Corrosion limited by transport of an oxidant

When a corrosive environment contains only small concentrations of an oxidizing agent, the corrosion rate is often limited by its transport. This behavior is typically found for corrosion by oxygen in neutral media or corrosion by protons in weakly acidic environments.

Corrosion by oxygen

The ***solubility of oxygen*** in water is relatively low. According to ***Henry's law***, it depends on the partial pressure:

$$c_{O_2} = k_H P_{O_2} \tag{4.152}$$

At 25 °C, the Henry coefficient is equal to $k_H = 1.30 \times 10^{-3}$ mol dm^{-3} bar^{-1}. This corresponds to an equilibrium oxygen concentration of about 2.5×10^{-4} mol l^{-1} (8 mg l^{-1}) in water at equilibrium with air ($P_{O2} \approx 0.2$ bar). At higher temperatures the solubility of oxygen is even smaller.

Figure 4.36 shows the solubility of oxygen in contact with air at a total pressure of $P = P_{air} + P_{H2O} = 1$ bar. The concentration of dissolved oxygen decreases with temperature and becomes zero at 100 °C. The observed behavior is the result of two phenomena: the decrease in the Henry coefficient with increasing temperature and the increase of the partial pressure of water at the expense of that of oxygen. When T = 100 °C the partial pressure of water becomes equal to the atmospheric pressure, $P_{H2O} = 1$ bar, it follows that $P_{O2}= 0$, and the concentration of dissolved oxygen is zero. The second curve of Figure 4.36 shows the solubility at a constant oxygen partial pressure, $P_{O2} = 1$ bar.

© 2007, first edition, EPFL Press

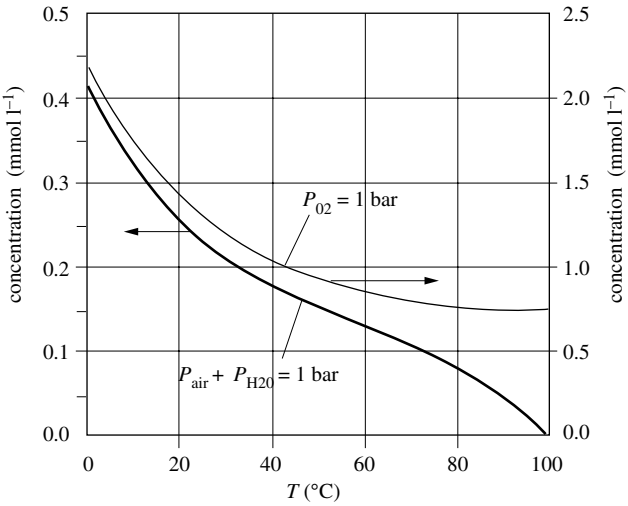


Figure 4.36 Variation of oxygen solubility in water as a function of the temperature for a total constant pressure of $P = P_{O2} + P_{N2} + P_{H2O} = 1$ bar and for a constant oxygen partial pressure of $P_{O2} = 1$ bar.

The low solubility of oxygen in water, it is even lower in salt solutions, limits the oxygen flux at the metal that governs the rate of corrosion. Figure 4.37 shows, in the form of an Evans diagram, the partial anodic and cathodic current densities during the corrosion of a metal by oxygen. It is assumed that the dissolution kinetics follow the Butler-Volmer equation and that the corrosion potential is located within the region of the limiting current plateau for oxygen reduction. It follows from the figure that the corrosion current density, i_{cor}, in this case is equal to the limiting current density for oxygen reduction $i_{l,O2}$ and therefore it varies with the thickness δ of the diffusion layer which is controlled by hydrodynamic conditions.

$$i_{cor} = -i_{l,O_2} = -4\,F\,D_{O_2}\,\frac{c_{O_2,b}}{\delta} \tag{4.153}$$

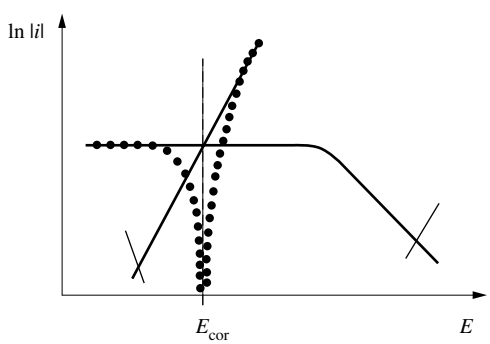


Figure 4.37 Evans diagram for corrosion limited by oxygen transport.

© 2007, first edition, EPFL Press

An increase in temperature produces two opposing effects: on one hand, the solubility of oxygen decreases, and, on the other hand, the diffusion coefficient increases. At constant total pressure, the limiting current density therefore goes through a maximum as a function of temperature [14,15]. Near the boiling point, its value becomes very small.

Corrosion of iron in neutral, aerated media

Figure 4.38 shows a series of cathodic polarization curves, measured in an aerated solution of 0.5M NaCl, on a rotating disk electrode made of iron [14]. The dashed lines represent the partial current densities. Cathodic current plateaus are observed that represent the limiting current for oxygen reduction. Their height varies with the electrode rotation rate. Figure 4.39 shows the corrosion current density i_{cor} of an iron electrode, plotted as a function of the square root of the rotation rate of the disk electrode. The experiments were carried out in a solution of 0.2M Na_2SO_4 of pH 2.7 using chemical analysis of Fe^{2+}ions for the determination of the corrosion rate. Also shown in the figure is the limiting current density for the reduction of oxygen determined by polarization measurements. Both quantities vary linearly with the square root of the rotation rate (Levich equation). This confirms that the rate of iron corrosion under the conditions of the experiment was indeed limited by oxygen transport. Figures 4.38 and 4.39 show comparable limiting current densities, indicating a similar oxygen solubility in the two electrolytes.

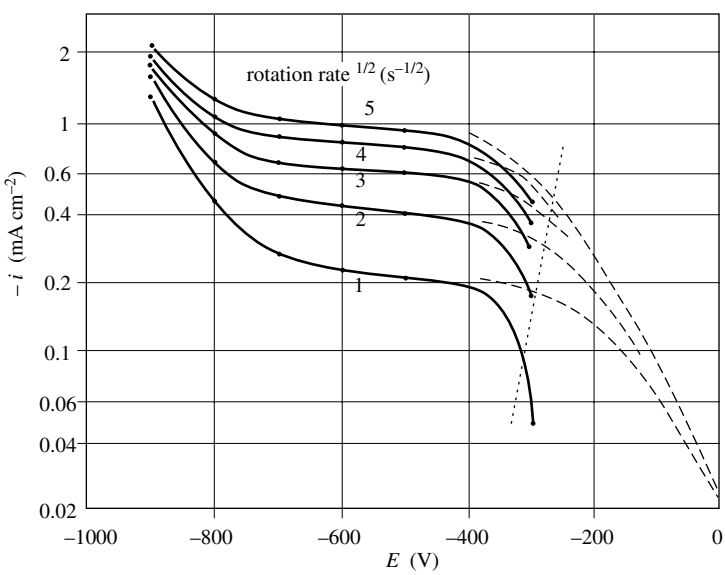


Figure 4.38 Polarization curves for Fe in aerated 0.5M NaCl measured with a rotating disk electrode at 20 °C by applying different rotation rates [14]. The dashed lines show the anodic and cathodic partial current densities.

© 2007, first edition, EPFL Press

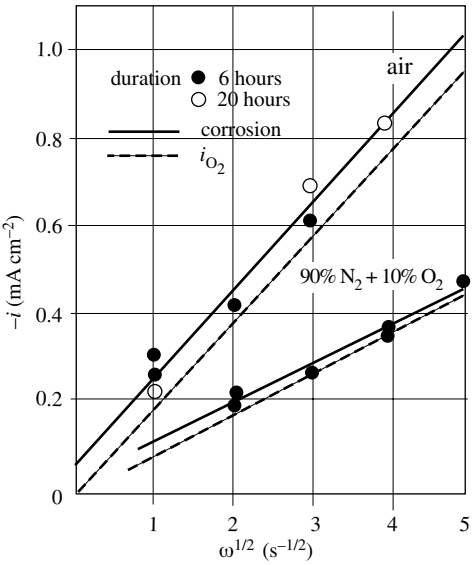


Figure 4.39 Variation of corrosion rate and of cathodic limiting current with rotation rate in air and in N_2 +10% O_2. The rates after 6 and 20 hours are shown. Dashed line: limiting current density of the oxygen reduction measured under comparable conditions. Electrolyte: 0.2M Na_2SO_4, pH 2.7.

The influence of hydrodynamic conditions on the corrosion rate is not always straightforward. The reason is that often the surface of metals exposed to a corrosive environment becomes covered with corrosion products which impede oxygen access to the surface. As they form, the corrosion rate therefore gradually decreases with time. A typical example of this behavior is the formation of carbonate or sulfate scales on steel exposed to natural waters.

Figure 4.40 shows the quantity of metal lost due to corrosion as a function of exposure time from a steel pipe filled with water containing sulfate ions. The mass loss is not a linear function of time because the corrosion rate decreases due to the formation of scale on the pipe wall. On the other hand, the higher corrosion rate observed at higher Reynolds number indicates that mass transport plays a critical role.

The observed behavior can be interpreted by a simple model that takes into account the mass transport of oxygen in the diffusion layer and in a porous layer formed by the corrosion products. The following equation gives the overall mass transport rate of oxygen:

$$N_{O_2} = k_{mf}\ c_{O_2,b} \tag{4.154}$$

Here, N_{O2} designates the oxygen flux at the surface and $c_{O2,b}$ the oxygen concentration in the bulk. The overall mass transport coefficient, k_{mf}, contains two terms: the coefficient k_m which characterizes the convective transport and the coefficient k_f,

© 2007, first edition, EPFL Press

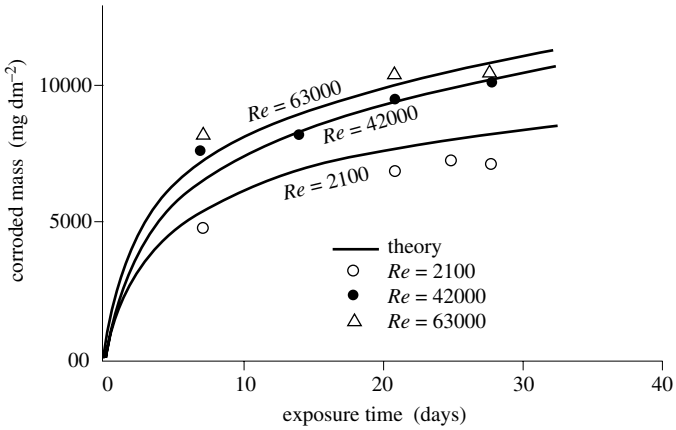


Figure 4.40 Mass loss of a pipe wall due to corrosion at different flow rates of natural water containing 270 mg l^{-1} SO_4^{2-} under conditions of forced convection for different Reynold's numbers [16].

which characterizes the rate of transport in the porous film (θ_p = porosity, δ_f = thickness of the film).

$$\frac{1}{k_{mf}} = \frac{1}{k_m} + \frac{1}{k_f} \tag{4.155}$$

$$k_m = \frac{D_{O_2}}{\delta}$$

$$k_f = \frac{D_{O_2}\theta_p}{\delta_f}$$

Equation (4.155) shows that the smaller of the two coefficients (k_m or k_f) limits the corrosion rate. Generally, when a solid film (carbonates, oxides, sulfates, etc.) forms on a metal surface, the corrosion rate progressively decreases with time. While this is a fortunate phenomenon it also complicates efforts to develop accelerated corrosion tests. Indeed, in Figure 4.40, if one would extrapolate the results obtained after two or three days only, one would grossly overestimate the rate of corrosion that determines the lifetime of the steel pipe.

Corrosion in weakly acidic environments

In acid solutions, the reduction of protons is the main cathodic reaction:

$$H^+ + e = \frac{1}{2} H_2 \tag{4.156}$$

© 2007, first edition, EPFL Press

At high acid concentration, the reaction rate of protons is normally limited by charge transfer at the electrode-electrolyte interface (Sect. 4.2). In weakly acidic environments, however, the reaction rate may be limited by the rate of transport. The hydrodynamic conditions and the pH then determine the value of the limiting current. Neglecting migration effects its value is given by:

$$i_l = -F\,D_{H^+}\frac{c_{H^+,b}}{\delta} \tag{4.157}$$

The width of the limiting-current plateau is limited by the onset of hydrogen formation from the reduction of water molecules:

$$H_2O + e = \frac{1}{2}H_2 + OH^- \tag{4.158}$$

Figure 4.41 presents a series of cathodic polarization curves measured on iron [17] in de-aerated 4% NaCl solutions of different pH. The limiting current is highest at the lowest pH values, because the cathodic current is limited by proton transport. If, in such a system, the corrosion potential lies in the plateau region, the corrosion rate is entirely limited by the rate of mass transport of protons. Figure 4.42 shows the effect of HCl concentration on the corrosion rate of a magnesium cylinder rotating at different speeds. At concentrations under 0.1M, the corrosion rate varies with the

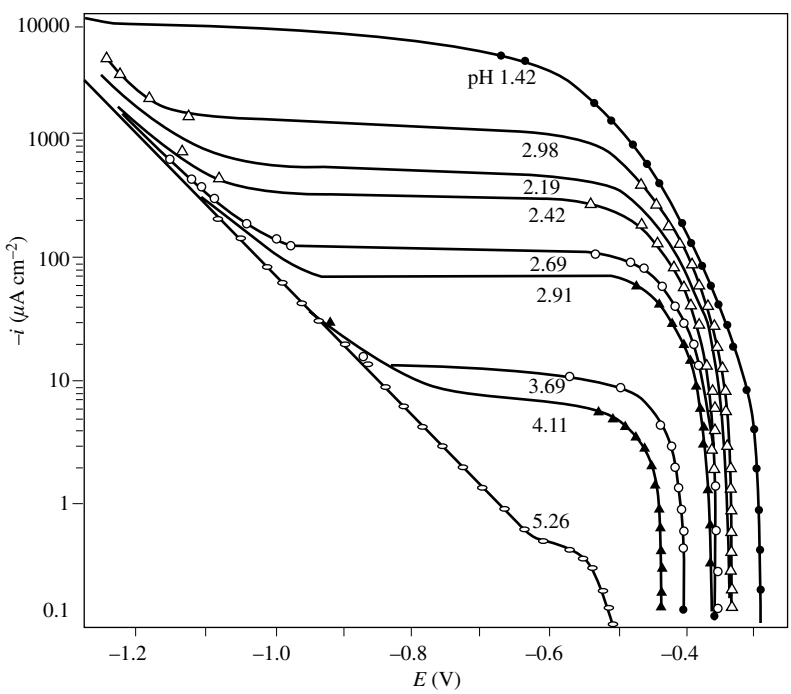


Figure 4.41 Cathodic polarization curves of iron in deaerated 4% NaCl at different pH [17].

© 2007, first edition, EPFL Press

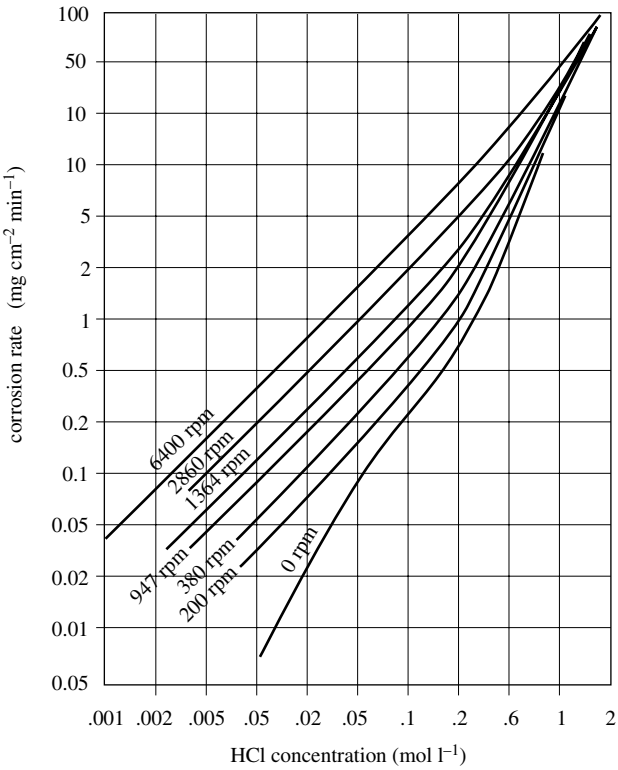


Figure 4.42 Corrosion rate of a magnesium rotating cylinder (11-mm diameter) as a function of HCl concentration at different rotation rates. Deaerated solutions at T= 25 °C [18].

rotation rate in agreement with the relation indicated in Table 4.27. This confirms that the transport of the H^+ ions limits the rate of magnesium corrosion under these conditions. At sufficiently high HCl concentration, the corrosion rate becomes independent of the rotation rate, because now it is controlled by charge transfer rather than by mass transport.

4.3.5 Corrosion limited by the transport of reaction products

Maximum rate of dissolution

Whenever a metal dissolves, a metal ion concentration gradient is established at the surface (Figure 4.20). The local metal ion concentration at the surface increases with the rate of dissolution, but under steady state conditions it does not exceed the saturation concentration, because precipitation will occur leading to formation of a ***salt film***. The maximum dissolution rate therefore is reached when the concentration at the surface becomes saturated with respect to the reaction products. The mass transport of metal ions from the surface of the film into the bulk electrolyte then limits the metal dissolution rate.

© 2007, first edition, EPFL Press

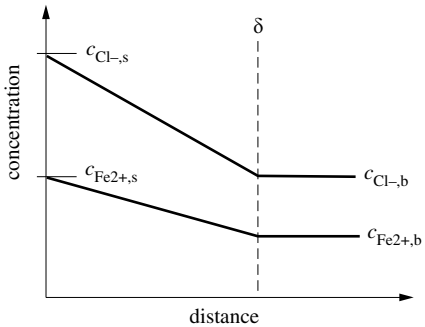


Figure 4.43 Concentration profiles at the electrode surface during anodic dissolution of Fe into a $FeCl_2$ solution.

The anodic behavior of iron in a binary electrolyte of $FeCl_2$ illustrates this phenomenon. In the absence of oxygen, the dissolution of iron produces ferrous ions.

$$Fe \rightarrow Fe^{2+} + 2e \tag{4.159}$$

The concentration profile of Fe^{2+} at the anode is given by equation (4.143), integrated between the limits $y = 0$, $c_{Fe2+} = c_{Fe2+,s}$ and $y = \delta$, $c_{Fe2+} = c_{Fe2+,b}$ where the subscripts (s) and (b) indicate the concentrations at the surface and in the bulk, respectively.

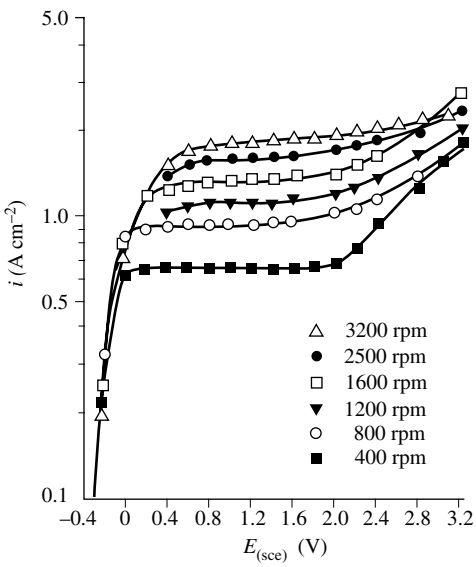


Figure 4.44 Anodic polarization curves for iron in 3M $FeCl_2$ solution at different rotation rates of a rotating disk electrode [19].

© 2007, first edition, EPFL Press

$$i = nF\left(1 - \frac{z_+}{z_-}\right) D_+ \frac{c_{Fe^{2+},s} - c_{Fe^{2+},b}}{\delta} \tag{4.160}$$

At the limiting current $c_{Fe2+,s} = c_{Fe2+,sat}$:

$$i_l = nF\left(1 - \frac{z_+}{z_-}\right) D_+ \frac{c_{Fe^{2+},sat} - c_{Fe^{2+},b}}{\delta} \tag{4.161}$$

In this case $n = 2$, $z_+ = 2$ and $z_- = -1$. Figure 4.43 schematically shows the concentration profile of Fe^{2+} and Cl^- ions. The electroneutrality condition requires that the chloride concentration at each point in the diffusion layer is equal to twice the ferrous ion concentration. At the limiting current, one finds thus for the surface concentration:

$$c_{Fe^{2+},s} = \frac{1}{2} c_{Cl^-,s} = c_{FeCl_2,sat} \tag{4.162}$$

where the subscript *sat* refers to the saturation concentration.

According to equation (4.161), the anodic limiting current density measured on a rotating disk electrode varies linearly with the square root of the rotation rate. In addition, it is proportional to the concentration difference $c_{Fe2+,sat,} - c_{Fe2+,b}$. Experimental results confirm this behavior. Figure 4.44 shows the anodic polarization curves obtained on a rotating disk electrode in a 3M $FeCl_2$ solution at different rotation rates. A well-defined current plateau extends across a region of up to two

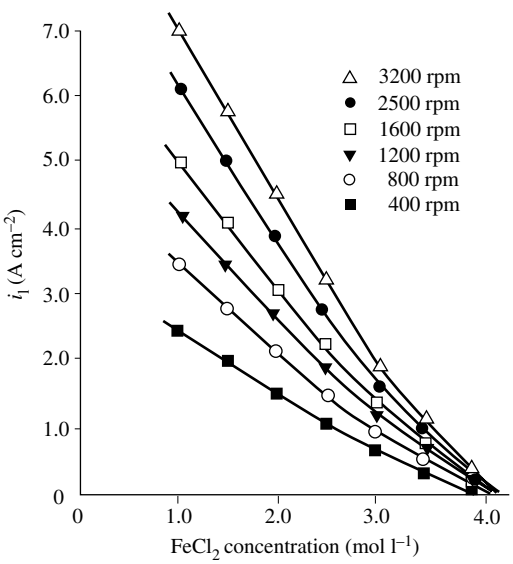


Figure 4.45 Anodic dissolution of Fe in $FeCl_2$: Variation of anodic limiting current density with bulk concentraion for different rotation rates of a rotating disk electrode [19].

© 2007, first edition, EPFL Press

volts. The value of the limiting current increases proportionately to $\omega^{1/2}$ (not shown) and it decreases as the concentration $FeCl_2$ in the electrolyte increases (Figure 4.45) in agreement with equation (4.161).

When a metal M dissolves in an electrolyte that does not contain its ions, for example when iron dissolves in a NaCl solution, the bulk metal ion concentration is zero, $c_{Mn+,b} \approx 0$. The limiting current density in the presence of a salt film can then be estimated by the following equation:

$$i_1 = n\,F\,D_{eff}\,\frac{c_{Mn+,sat}}{\delta} \tag{4.163}$$

where D_{eff} represents the effective diffusion coefficient as defined by (4.151), and $c_{Mn+,sat}$ is the saturation concentration of the salt that precipitates at the anode. Its value is usually affected the concentrations of the other salts present in the electrolyte and, in certain cases, by the pH, because of the possible precipitation of hydroxides.

Corrosion rate

If the corrosion potential of a metal is located within the plateau of the anodic limiting current (Figure 4.46), the corrosion rate equals the anodic limiting current density. It is therefore limited by the transport of reaction products from the surface of the metal to the bulk electrolyte. Its value can be estimated from equation (4.163).

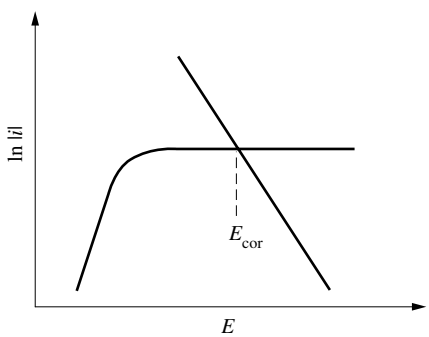


Figure 4.46 Evans diagram for corrosion controlled by mass transport of anodic reaction products.

Figure 4.47 shows the variation of the corrosion rate of a steel sample in sulfuric acid ($T = 60°$ C, 68% H_2SO_4), by plotting the dimensionless Sherwood number as a function of the Reynolds number. The corrosion rate was determined by gravimetric analysis. The data shown fit well with the prediction of the dimensionless correlation $Sh = f(Re, Sc)$ for the hydrodynamic conditions of the experiment. The rate of corrosion therefore is controlled by mass transport. With Figure 4.48, it is possible to identify the limiting species. It presents the results of laboratory tests, carried out with a rotating cylinder made of steel (69% H_2SO_4, 40 °C, 900 RPM), in solutions of different concentrations of Fe^{2+}. The corrosion rate decreases with increasing Fe^{2+}concentration as predicted by the model (4.161). It is therefore limited by the

© 2007, first edition, EPFL Press

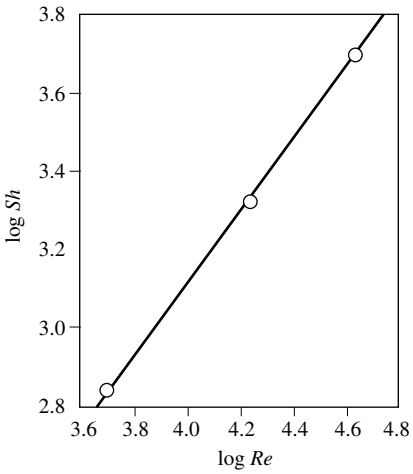


Figure 4.47 Corrosion controlled by mass transport of anodic reaction products: variation of the dimensionless corrosion rate (Sherwood number) as a function of Reynolds number for corrosion of a steel pipe in 68% H_2SO_4 at 60 °C [20].

transport of Fe^{2+} ions from the steel surface to the bulk solution. As sulfuric acid becomes more concentrated, the solubility of $FeSO_4$ decreases. Normal steel, which corrodes rapidly in solutions of sulfuric acid that are not too concentrated, resists well in highly concentrated sulfuric acid (≥ 96%), because the low solubility of $FeSO_4$ limits the flux of the ferrous ions and, therefore, the rate of corrosion. In the same way, the

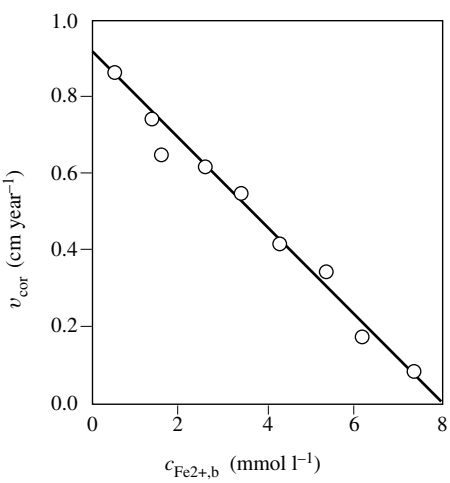


Figure 4.48 Corrosion controlled by mass transport of anodic reaction products: corrosion rate of steel in 68% H_2SO_4 at 60 °C as a function of Fe^{2+} bulk concentration. The rotation rate is 900 rpm, and the cylinder diameter is 1.9 cm [20].

© 2007, first edition, EPFL Press

low solubility of $PbSO_4$ explains why lead resists well to corrosive attack in sulfuric acid.

Electrode potentials at the limiting current

When a dissolved species reacts at an electrode under limiting current conditions its surface concentration is essentially zero. As a consequence, the value of the concentration overvoltage tends towards infinity and a current plateau is observed (equation 4.89). The situation is different when the limiting current is controlled by the transport of reaction products away from the electrode; the concentration overvoltage in this case remains generally small. Let us consider the example of Figure 4.45. For a $FeCl_2$ saturation concentration at the surface of about 4.2 M and a $FeCl_2$ bulk concentration of 3 M the concentration overvoltage calculated from equation (4.85) is on the order of 4 mV only. Clearly, concentration overvoltage can not explain the current plateaus shown in Figure 4.44, which extend over as much as 2V.

At the limiting current the surface concentration of reaction products reaches saturation and a salt film precipitates. The potential drop at the electrode is then determined essentially by the conduction properties of the surface film. For the case of iron dissolution in concentrated chloride solution the salt film was found to have a duplex structure, consisting of an inner compact and an outer porous layer [21]. The thickness of the compact layer, where most of the potential drop occurs, increases linearly with applied potential. On the other hand, the rate of dissolution of the salt layer is governed by mass transport and therefore is independent of potential. Thus the anodic current remains constant even if the potential difference across the salt film increases. Eventually, at sufficiently high potentials, other reaction phenomena may occur that lead to a renewed current rise, in a similar way as discussed in Section 4.3.1 for the limiting current of copper deposition.

General Bibliography

K. J. Vetter, *Electrochemical Kinetics*, Academic Press, New York (1967) 789 pp.
H. Kaesche, *Corrosion of Metals*, Springer Verlag, Berlin (2003) 594 pp.
J. S. Newman, *Electrochemical Systems*, 2nd edition, Prentice Hall, Englewood Cliffs (1991) 560 pp.

References

[1] C. Wagner, W. Traud, *Z. Elektrochemie* 44, 391 (1938).
[2] M. Stern, A. L. Geary, *J. Electrochem. Soc.* 104, 56 (1957).
[3] Norm ISO 8407 (1991).
[4] W. H. Ailor, *Handbook on Corrosion Testing and Evaluation,* The Electrochemical Society Inc., Princeton N.J. (1971) p. 136.
[5] Landolt-Börnstein, *Zahlenwerte und Funktionen*, 6th edition, Teil 5a, Springer Verlag Berlin (1969) p. 611, Teil 7 II (1960) p. 257.

© 2007, first edition, EPFL Press

[6] A. Bonnel, F. Dabosi, C. Delouis, M. Duprat, M. Keddam, B. Tribollet, *J. Electrochem. Soc.* 130, 753 (1983).
[7] J. S. Newman, *Electrochemical Systems,* 2nd edition, Prentice Hall, Englewood Cliffs (1991) p. 339.
[8] V. G. Levich, *Physicochemical Hydrodynamics*, Prentice Hall, Englewood Cliffs N.J. (1962) pp. 1 – 700.
[9] J. R. Selman, C. W. Tobias, *Adv. Chem. Engr.,* T. A. Drew, ed., vol. 10, Academic Press (1978) p. 211.
[10] N. Ibl, O. Dossenbach, *Comprehensive Treatise of Electrochemistry*, vol. 6, Plenum Press, New York (1982) p. 133.
[11] J. Newman *J. Electrochem. Soc.* 113, 501 (1966).
[12] K. J. Kretchner, C. H. Hammann, B. Fassbinder, *Electroanal. and Interf. Chem.* 60, 231 (1975).
[13] *Gmelins Handbuch der anorg. Chemie*, 8th edition, Lieferung 3, Verlag Chemie, Weinheim (1958) p. 460.
[14] K. Bohnenkamp, *Arch. Eisenhüttenwesen* 47, 253 (1976).
[15] H. H. Uhlig, R. W. Revie, *Corrosion and Corrosion Control,* 3rd edition, John Wiley, New York (1985) p. 95.
[16] L. W. Shemilt, C. Y. Cha, E. Fiadzigbe, A. B. Ponter, *Corr. Sci.* 20, 443 (1980).
[17] M. Stern, *J. Electrochem. Soc.* 102, 610 (1955).
[18] B. Roald, W. Beck, *J. Electrochem. Soc.* 98, 277 (1951).
[19] H. C. Kuo, D. Landolt, *Electrochim. Acta* 20, 393 (1975).
[20] B. T. Ellison, W. R. Schmeal, *J. Electrochem. Soc.* 125, 524 (1978).
[21] R. D. Grimm, A. C. West, D. Landolt, *J. Electrochem. Soc.* 139, 1622 (1992).

© 2007, first edition, EPFL Press

CHAPTER 5

EXPERIMENTAL STUDY OF ELECTRODE REACTIONS

5.1 ELECTROCHEMICAL KINETICS AND REACTION MECHANISMS

5.1.1 Standard rate constant of an electrode reaction

Consider an electrode reaction between two dissolved species involving the transfer of n electrons:

$$B_{\text{ox}} + n\,e = B_{\text{red}} \tag{5.1}$$

The corresponding Butler-Volmer equation (4.26) can be written as follows:

$$i = i_{\text{a}} + i_{\text{c}} = n\,F\,k_{\text{a}}c_{\text{red}}\exp\left(\frac{\alpha_{\text{a}}n\,F}{RT}E\right) - n\,F\,k_{\text{c}}c_{\text{ox}}\exp\left(-\frac{\alpha_{\text{c}}n\,F}{RT}E\right) \tag{5.2}$$

where in order to simplify, we set: $\alpha = \alpha_{\text{a}}$, $1-\alpha = \alpha_{\text{c}}$ and $c_{\text{red,s}} = c_{\text{red}}$, $c_{\text{ox,s}} = c_{\text{ox}}$. The rate constants k_{a} and k_{c} depend on the reference electrode (Section 4.2). Most often, they refer the standard hydrogen electrode, but sometimes in electrochemical kinetics it is preferable to refer the rate constants to the standard potential of the electrode reaction itself. We must then modify equation (5.2) in the following manner:

$$i = i_{\text{a}} + i_{\text{c}} = n\,F\,k_{\text{a}}^{0}c_{\text{red}}\exp\left(\frac{\alpha_{\text{a}}n\,F}{RT}\left(E-E^{0}\right)\right) - n\,F\,k_{\text{c}}^{0}c_{\text{ox}}\exp\left(-\frac{\alpha_{\text{c}}n\,F}{RT}\left(E-E^{0}\right)\right) \tag{5.3}$$

with:

$$k_{\text{a}}^{0} = k_{\text{a}}\exp\left(\frac{\alpha_{\text{a}}n\,F}{RT}E^{0}\right)$$

$$k_{\text{c}}^{0} = k_{\text{c}}\exp\left(\frac{\alpha_{\text{c}}n\,F}{RT}E^{0}\right)$$

© 2007, first edition, EPFL Press

At equilibrium, $E = E_{rev}$, $i = 0$:

$$n F k_a^0 c_{red} \exp\left(\frac{\alpha_a n F}{RT}\left(E_{rev} - E^0\right)\right) = n F k_c^0 c_{ox} \exp\left(-\frac{\alpha_c n F}{RT}\left(E_{rev} - E^0\right)\right) \quad (5.4)$$

By taking the logarithm and applying the relation $\alpha_a + \alpha_c = 1$, we obtain:

$$E_{rev} = E^0 + \frac{RT}{nF} \ln \frac{k_c^0}{k_a^0} + \frac{RT}{nF} \ln \frac{c_{ox}}{c_{red}} \quad (5.5)$$

For $k_a^0 = k_c^0 = k^0$ the expression (5.5) corresponds to the Nernst equation in which the activity has been replaced by concentration. The ***standard rate constant*** k^0 of an electrode reaction unambiguously characterizes its reaction rate. Equation (5.3) then becomes:

$$i = n F k^0 c_{red} \exp\left(\frac{\alpha_a n F}{RT}\left(E - E^0\right)\right) - n F k^0 c_{ox} \exp\left(-\frac{\alpha_c n F}{RT}\left(E - E^0\right)\right) \quad (5.6)$$

We can determine k^0 experimentally by an extrapolation of either the anodic or cathodic Tafel line to the standard potential. A second method to obtain k^0 is based on the measurement of the exchange current density:

$$i_0 = n F k^0 c_{red} \exp\left(\frac{\alpha_a n F}{RT}\left(E_{rev} - E^0\right)\right) = n F k^0 c_{ox} \exp\left(-\frac{\alpha_c n F}{RT}\left(E_{rev} - E^0\right)\right) \quad (5.7)$$

By setting

$$E_{rev} = E^0 + \frac{RT}{nF} \ln \frac{c_{ox}}{c_{red}} \quad (5.8)$$

we obtain (5.9).

$$i_0 = n F k^0 c_{red}^{(1-\alpha_a)} c_{ox}^{\alpha_a} = n F k^0 c_{ox}^{(1-\alpha_c)} c_{red}^{\alpha_c} \quad (5.9)$$

The charge transfer coefficients α_a and α_c are obtained from the variation of i_0 as a function of the concentrations c_{red} and c_{ox}.

$$\left.\frac{\partial \ln i_0}{\partial \ln c_{red}}\right|_{c_{ox}} = 1 - \alpha_a = \alpha_c \quad (5.10)$$

$$\left.\frac{\partial \ln i_0}{\partial \ln c_{ox}}\right|_{c_{red}} = \alpha_a = 1 - \alpha_c \quad (5.11)$$

© 2007, first edition, EPFL Press

5.1.2 Tafel coefficients for multi-step reactions

Most electrode reactions encountered in the field of corrosion involve the transfer of more than one electron. Such reactions take place in steps, of which the slowest, called the ***rate-determining step,*** abbreviated *RDS*, determines the overall reaction rate. In simple cases, one can identify the rate-determining step by an analysis of the measured Tafel slopes. In the so-called *quasi-equilibrium approach* one assumes that with the exception of the rate-limiting step, all other steps are at equilibrium. This greatly simplifies the mathematical equations for the reaction rate. More realistic approaches require numerical simulation and shall not be discussed here. To illustrate the quasi equilibrium approach to the study of multi-step electrode reactions we shall look at proposed mechanisms for the dissolution of copper and of iron.

Dissolution of copper

The dissolution of copper in sulfate solution corresponds to a two-step electrode reaction, the second step being rate determining.

$$\text{step I:} \qquad \mathrm{Cu} \rightarrow \mathrm{Cu}^{+} + e \tag{5.12}$$

$$\underline{\text{step II (RDS):} \qquad \mathrm{Cu}^{+} \rightarrow \mathrm{Cu}^{2+} + e} \tag{5.13}$$

$$\text{(I) + (II):} \qquad \mathrm{Cu} \rightarrow \mathrm{Cu}^{2+} + 2\,e \tag{5.14}$$

To obtain an analytical expression for the anodic Tafel coefficient we apply the Butler-Volmer equation to each step. Their respective rates are v_{I} and v_{II} [mol m^{-2} s^{-1}]. To simplify the equations, we use the abbreviation:

$$f = \frac{F}{RT} \tag{5.15}$$

We thus obtain:

$$v_{\mathrm{I}} = k_{\mathrm{a,I}} \exp\left(\alpha_{\mathrm{a,I}} f E\right) - k_{\mathrm{c,I}}\, c_{\mathrm{Cu^{+}}} \exp\left(-\alpha_{\mathrm{c,I}} f E\right) \tag{5.16}$$

$$v_{\mathrm{II}} = k_{\mathrm{a,II}}\, c_{\mathrm{Cu^{+}}} \exp\left(\alpha_{\mathrm{a,II}} f E\right) - k_{\mathrm{c,II}}\, c_{\mathrm{Cu^{2+}}} \exp\left(-\alpha_{\mathrm{c,II}} f E\right) \tag{5.17}$$

If step I is at equilibrium:

$$c_{\mathrm{Cu^{+}}} = \frac{k_{\mathrm{a,I}}}{k_{\mathrm{c,I}}} \exp\left(\alpha_{\mathrm{a,I}} f E\right) \exp\left(\alpha_{\mathrm{c,I}} f E\right) \tag{5.18}$$

Since by definition $\alpha_{\mathrm{a}} + \alpha_{\mathrm{c}} = 1$ and by setting $K_{\mathrm{I}} = k_{\mathrm{a,I}}/k_{\mathrm{c,I}}$, this relation becomes:

$$c_{\mathrm{Cu^{+}}} = K_{\mathrm{I}} \exp\left(f E\right) \tag{5.19}$$

© 2007, first edition, EPFL Press

Because the rate determining second step is far from equilibrium we can neglect in (5.17) the cathodic partial reaction. With (5.18) this leads to the following expression for the anodic partial reaction:

$$v_{\mathrm{a,II}} = k_{\mathrm{a,II}} K_{\mathrm{I}} \exp\left(f E\right) \exp\left(\alpha_{\mathrm{a,II}} f E\right) \tag{5.20}$$

which is equivalent to

$$v_{\mathrm{a,II}} = k_{\mathrm{a,II}} K_{\mathrm{I}} \exp\left(\left(1+\alpha_{\mathrm{a,II}}\right) f E\right) \tag{5.21}$$

According to Faraday's law, the current density is equal to

$$i = 2F\left(v_{\mathrm{a}} - v_{\mathrm{c}}\right) \tag{5.22}$$

Under steady-state conditions the reaction rate is the same for each step and corresponds to the reaction rate of the overall electrode reaction: $v = v_{\mathrm{I}} = v_{\mathrm{II}}$. In the anodic Tafel region, $v_{\mathrm{c}} \approx 0$, and thus we find for the current:

$$i = i_{\mathrm{a}} = 2F k_{\mathrm{a,II}} K_{\mathrm{I}} \exp\left(\left(1+\alpha_{\mathrm{a,II}}\right) f E\right) \tag{5.23}$$

This allows us to obtain the anodic Tafel coefficient.

$$\beta_{\mathrm{a}} = \frac{\mathrm{d}E}{\mathrm{d}\ln i_{\mathrm{a}}} = \frac{RT}{\left(1+\alpha_{\mathrm{a,II}}\right)F} \tag{5.24}$$

At 25 °C, and assuming that the anodic charge transfer coefficient has the value $\alpha_{\mathrm{a,II}} = \alpha_{\mathrm{a}} = 0.5$, this yields $\beta_{\mathrm{a}} = 17$ mV.

By analogous reasoning, the cathodic Tafel coefficient is given by the relation:

$$\beta_{\mathrm{c}} = -\frac{\mathrm{d}E}{\mathrm{d}\ln\left|i_{\mathrm{c}}\right|} = \frac{RT}{\left(1-\alpha_{\mathrm{c,II}}\right)F} \tag{5.25}$$

With the values $\alpha_{\mathrm{c,II}} = \alpha_{\mathrm{c}} = 0.5$ this leads to $\beta_{\mathrm{c}} = 52$ mV. The calculated values of β_{a} and β_{c} correspond to anodic and cathodic Tafel slopes, respectively, of $b_{\mathrm{a}} = 2.3\beta_{\mathrm{a}} = 40$ mV and $b_{\mathrm{c}} = 2.3\beta_{\mathrm{c}} = 120$ mV. The polarization data shown in Figure 5.1 confirm these values and hence are consistent with the proposed mechanism.

The preceding example shows how a measurement of anodic and cathodic Tafel slopes can be used to identify an electrode reaction mechanism. For this one analyzes possible mechanisms and compares the predicted Tafel coefficients with those observed experimentally. Unfortunately, this approach has several shortcomings. First of all, it does not always lead to unambiguous results. Especially, for electrode reactions involving several steps more than one reaction mechanism may lead to a given set of values for β_{a} and β_{c} . Furthermore, the presented analysis includes the arbitrary hypothesis that the value of the transfer coefficients α_{a} and α_{c} is equal to 0.5, indicating a symmetrical activation energy barrier. Finally, the quasi equilibrium hypothesis is valid only if the rate of the rate determining step differs greatly from that of the preceding or following steps over the whole potential range studied.

© 2007, first edition, EPFL Press

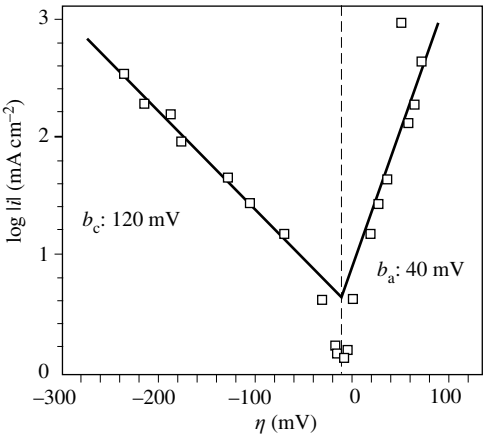


Figure 5.1 Polarization curve of copper in 0.5 M H_2SO_4 + 0.075 M $CuSO_4$.

Dissolution of iron

In weakly acid environments, the dissolution rate of iron varies with the pH. Different mechanisms have been proposed to explain this behavior, the best known being those attributed to Bockris and to Heusler [2].

The Bockris mechanism invokes a three-step dissolution, the step (II) being rate determining:

$$\text{step I:} \quad Fe + OH^- \rightarrow FeOH_{ads} + e \tag{5.26}$$

$$\text{step II (RDS):} \quad FeOH_{ads} \rightarrow FeOH^+ + e \tag{5.27}$$

$$\underline{\text{step III:} \quad FeOH^+ + H^+ \rightarrow Fe^{2+} + H_2O} \tag{5.28}$$

$$\text{(I)+(II)+(III):} \quad Fe \rightarrow Fe^{2+} + 2\,e \tag{5.29}$$

The Heusler mechanism can be written as follows:

$$\text{step I:} \quad Fe + OH^- \rightarrow FeOH_{ads} + e \tag{5.26}$$

$$\text{step II (RDS):} \quad Fe + OH^- + FeOH_{ads} \rightarrow FeOH^+ + FeOH_{ads} + 2e \tag{5.30}$$

$$\underline{\text{step III:} \quad FeOH^+ + H^+ \rightarrow Fe^{2+} + H_2O} \tag{5.28}$$

$$\text{(II)+(III):} \quad Fe \rightarrow Fe^{2+} + 2\,e \tag{5.29}$$

© 2007, first edition, EPFL Press

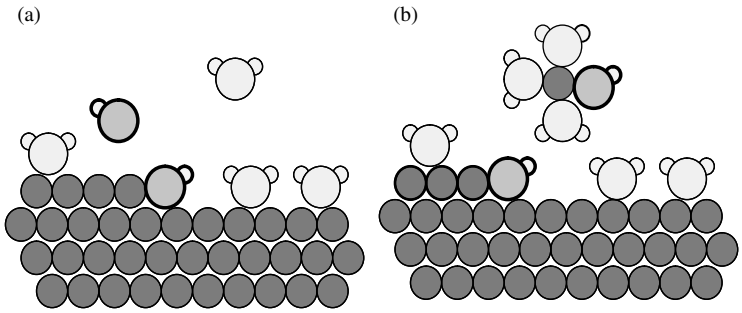


Figure 5.2 Dissolution of iron according to catalytic (Heusler) mechanism : (a) adsorption of OH^- at a kink site, (b) dissolution of a $FeOH^+$ species with regeneration of $FeOH_{ads}$.

This mechanism assumes the simultaneous transfer of two electrons in the rate-determining step (II), which takes place on catalytic sites that have been formed in a first step. In the above equations, the catalytic sites correspond to the species $FeOH_{ads}$. Physically, these sites are kink sites, whose quantity depends on the potential (step I). These positions are regenerated whenever an iron atom leaves in the form of the hydrated $FeOH^+$ ion. In this way, step (II) does not affect the number of catalytic sites (see Figure 5.2). Because the concentration of catalytic sites remains constant, the first step does not contribute to the current under steady state conditions. The third step indicates that the dissolved ions $FeOH^+$ are in equilibrium with hydrated Fe^{2+}.

Both mechanisms described above have in common that OH^- ions participate in the charge-transfer reaction. This leads to the formation of $FeOH_{ads}$ and $FeOH^+$ as reaction intermediates and to a pH dependence of the dissolution rate.

In order to calculate the anodic Tafel coefficient for the ***Bockris mechanism***, we use the quasi equilibrium approach and write the rate expression for the rate-determining step (II):

$$v_{a,\,II} = k_{a,\,II}\,\theta_{FeOH(ads)}\,\exp\left(\alpha_a\,f\,E\right) \tag{5.31}$$

The equilibrium (I) allows us to express the surface coverage of the adsorbed species $FeOH_{ads}$, $\theta_{FeOH(ads)}$, as a function of the potential.

$$k_{a,\,I}\,c_{OH^-}\,\exp\left(\alpha_{a,\,I}\,f\,E\right) = k_{c,\,I}\,\theta_{FeOH(ads)}\,\exp\left(-\alpha_{c,\,I}\,f\,E\right) \tag{5.32}$$

By setting $\alpha_{a,1} + \alpha_{c,1} = \alpha_a + \alpha_c = 1$ and $k_{a,1}/k_{c,1} = K_1$, we obtain:

$$\theta_{FeOH(ads)} = K_I\,c_{OH^-}\,\exp\left(f\,E\right) \tag{5.33}$$

$$i_a = 2\,F\,v_{a,II} = 2\,F\,k_{a,\,II}\,K_I\,c_{OH^-}\,\exp\left(\left(1+\alpha_a\right)f\,E\right) \tag{5.34}$$

© 2007, first edition, EPFL Press

The anodic Tafel coefficient thus becomes:

$$\beta_a = \frac{RT}{(1+\alpha_a)F} \tag{5.35}$$

If $\alpha_a = 0.5$ and $T = 25$ °C, this results in $\beta_a = 17$ mV.

According to the ***Heusler mechanism***, the anodic current density of the rate-determining step is written as:

$$i_a = 2F\,v_{a,II} = 2F\,k_{a,II}\,c_{OH^-}\,\theta_{FeOH(ads)}\exp\left(2\alpha_{a,II}\,f\,E\right) \tag{5.36}$$

In the exponential term, the factor of two corresponds to the number of electrons transferred in this step. The quantity $\theta_{FeOH(ads)}$ represents the surface coverage of catalytic sites, which depends on the potential. If step (I) is at equilibrium, equation (5.33) is applicable. By introducing this expression into (5.36), we get:

$$i_a = 2F\,v_{a,II} = 2F\,k_{a,II}\,c_{OH^-}^2\exp\left(\frac{(1+2\alpha_{a,II})F}{RT}E\right) \tag{5.37}$$

The anodic Tafel coefficient therefore corresponds to:

$$\beta_a = \frac{RT}{(1+2\alpha_{a,II})F} \tag{5.38}$$

By assuming that $\alpha_{a,II} = \alpha_a = 0.5$, we obtain $\beta_a = 13$ mV. This value corresponds to a Tafel slope of $b_a = 30$ mV per decade, on the base-10 logarithmic scale. Figure 5.3

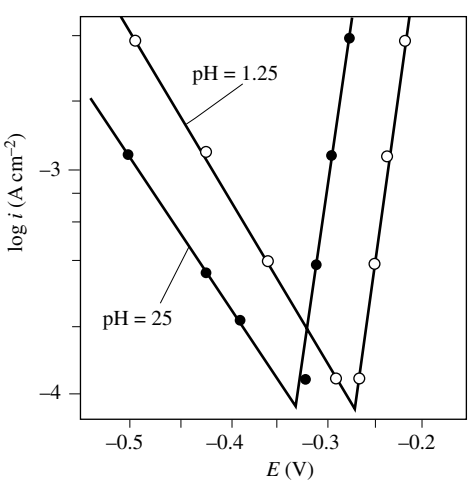


Figure 5.3 Polarization curves of Fe in 0.5 M $NaClO_4$ + $HClO_4$ of pH 1.25(O) and pH 2.5 (•) at 20 °C (adapted from ref. [2]).

© 2007, first edition, EPFL Press

shows the experimental results obtained in sulfate solutions at different pH [2]. The anodic Tafel slope of 30 mV is consistent with the described mechanism.

5.1.3 Mechanism of the hydrogen electrode

The hydrogen electrode has been extensively studied because of its importance in corrosion and in fuel cell technology. We shall look here at the kinetics of the reduction of protons, which determines the corrosion rate in acid environments (Chapter 4). The cathodic reaction involves two electrons:

$$2\,H^+ + 2\,e \rightarrow H_2 \tag{5.39}$$

The charge-transfer reaction at the metal-electrolyte interface proceeds in two steps. First a proton is reduced to an adsorbed hydrogen atom (*Volmer reaction*). In a second step, a hydrogen molecule forms, either by chemical reaction between two adsorbed hydrogen atoms (*Tafel reaction*) or via an electrochemical reaction between an adsorbed hydrogen atom and a proton (*Heyrovsky reaction*). Finally, the hydrogen molecule leaves the cathode, either in the form of gas bubbles or by diffusion in the electrolyte. Under certain conditions, the adsorbed hydrogen atoms can dissolve into the metal instead of forming hydrogen molecules (See Chapter 11).

Volmer-Heyrovsky Mechanism

According to this mechanism, the formation of a hydrogen molecule proceeds in two consecutive electrochemical steps.

$$\text{step I (RDS):} \quad H^+ + e \rightarrow H_{ads} \tag{5.40}$$

$$\underline{\text{step II:} \quad H^+ + H_{ads} + e \rightarrow H_2} \tag{5.41}$$

$$\text{(I) + (II):} \quad 2\,H^+ + 2\,e \rightarrow H_2 \tag{5.42}$$

Step (II) is a charge-transfer reaction, in which the adsorbed hydrogen atom reacts with a proton and with an electron provided by the electrode. If step (I), called *Volmer reaction*, is rate limiting and, if in addition, $\alpha_{I,a} = \alpha_{II,a} = \alpha_a$ and $\alpha_{I,c} = \alpha_{II,c} = \alpha_c$, the reaction rate, v_I, can be expressed as follows:

$$\begin{aligned} v_I &= v_{I,a} - v_{I,c} \\ &= k_{a,I}\theta_H \exp(\alpha_a f E) - k_{c,I}(1-\theta_H)\, c_{H^+} \exp(-\alpha_c f E) \end{aligned} \tag{5.43}$$

The rate of the anodic partial reaction, $v_{I,a}$, is proportional to the amount of adsorbed hydrogen, expressed by the coverage θ_H, which represents the fraction of adsorption sites occupied by hydrogen ($0 < \theta_H < 1$). The rate of the cathodic partial reaction $v_{I,c}$ is proportional to the fraction of non-occupied sites, $1-\theta_H$. The cathodic partial current density thus becomes:

© 2007, first edition, EPFL Press

$$i_c = -2Fv_{I,c} = -2Fk_{c,I}\left(1-\theta_H\right)c_{H^+}\exp\left(-\alpha_c f E\right) \tag{5.44}$$

This corresponds to a cathodic Tafel coefficient of

$$\beta_c = \frac{RT}{\alpha_c F} \tag{5.45}$$

For T = 25 °C and with α_c = 0.5, this yields a cathodic Tafel slope of β_c = 120 mV per decade. Figure 5.4 shows the experimental results for different metals in sulfuric acid solutions obtained by a number of research groups [3]. Most of the Tafel lines in fact exhibit a slope of β_c = 120 mV/decade.

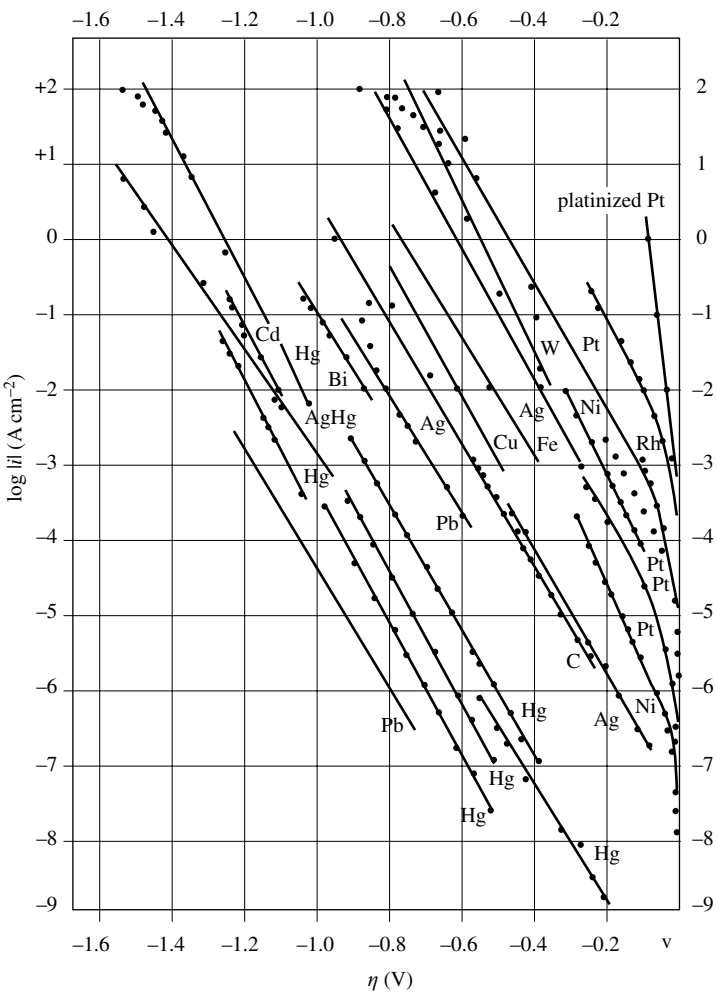


Figure 5.4 Cathodic polarization curves of hydrogen evolution on different metals in sulfuric acid (adapted from ref. [3]).

© 2007, first edition, EPFL Press

Volmer-Tafel mechanism

This mechanism differs from the preceding one in its second step, called *Tafel reaction*, which is chemical rather than electrochemical in nature. Furthermore, the second step is assumed to be rate-determining.

$$\text{step I:} \qquad H^+ + e \rightarrow H_{ads} \quad | \; 2\times \tag{5.40}$$

$$\underline{\text{step II (RDS):} \qquad H_{ads} + H_{ads} \rightarrow H_2} \tag{5.46}$$

$$\text{(I) + (II):} \qquad 2\,H^+ + 2\,e \rightarrow H_2 \tag{5.42}$$

If the first step would limit the reaction rate, the cathodic Tafel coefficient would be identical to that of the Volmer-Heyrovsky mechanism (the anodic Tafel coefficient would be different though). On the other hand, if the second step is rate determining the cathodic Tafel slope differs. The rate of the reaction determining step, v_{II} , is given by

$$v_{II} = k_{II,a}\left(1-\theta_H\right)c_{H_2} - K_{II,c}\left(\theta_H\right)^2 \tag{5.47}$$

and the rate of step I is equal to

$$v_I = k_{a,I}\theta_H \exp\left(\alpha_a f E\right) - k_{c,I}\left(1-\theta_H\right)c_{H^+} \exp\left(-\alpha_c f E\right) \tag{5.48}$$

Assuming that step (I) is at equilibrium:

$$k_{a,I}\theta_H \exp\left(\alpha_a f E\right) = k_{c,I}\left(1-\theta_H\right)c_{H^+} \exp\left(-\alpha_c f E\right) \tag{5.49}$$

$$\frac{\theta_H}{1-\theta_H} = \frac{k_{c,I}}{k_{a,I}} c_{H^+} \exp\left(-(\alpha_a+\alpha_c) f E\right) \tag{5.50}$$

According to (5.50) θ_H is an exponential function of the potential. For $\theta_H << 1$, i.e., if only a small fraction of the available adsorption sites of the metal surface are occupied, then $(1-\theta_H) \approx 1$. By setting $k_{c,1}/k_{a,1} = K'_I$, and $\alpha_a + \alpha_c = 1$, we obtain thus (5.51).

$$\theta_H = K'_I\, c_{H^+} \exp\left(-f E\right) \tag{5.51}$$

The cathodic partial current density is thus given by

$$i_c = 2F\, v_{II} = -2F\, k_{II,c}\left(\theta_H\right)^2 = -2F\, k_{II,c} K'^2_I\, c^2_{H^+} \exp\left(-2 f E\right) \tag{5.52}$$

According to this equation the cathodic Tafel coefficient is equal to

$$\beta_c = \frac{R\,T}{2\,F} \tag{5.53}$$

© 2007, first edition, EPFL Press

For 25 °C this yields a Tafel slope 30 mV per decade. In Figure 5.4, the results obtained for platinized platinum (a platinum electrode onto which a porous layer of platinum has been deposited as a catalyst) seem to agree with this mechanism.

Exchange-current density of the hydrogen electrode

The mechanisms described in this section provide a theoretical interpretation of the measured Tafel slopes, but they do not allow us to predict a value for the exchange-current density, which to a large extent determines the reaction rate. Its value depends on both the nature of the metal and electrolyte composition. Figure 5.5 shows the influence of the metal. The exchange-current density, measured on different metals in 1M acid solutions is presented as a function of the bond energy between the metal and hydrogen. Between the weakly catalytic metals (such as lead, zinc or mercury) and the strongly catalytic metals (such as platinum, rhodium, or palladium) the values of i_0 vary over 6 orders of magnitude, from 10^{-12} A cm^{-2} to 10^{-6} A cm^{-2}.

The curve presented in Figure 5.5 shows a maximum that can be explained in the following way. If the metal-hydrogen bond energy is too weak, the adsorption of hydrogen is not favored and, as a result, the reaction rate remains low. Conversely, a very high M-H bond energy increases the activation energy of the recombination reaction (step II); in this case, the hydrogen leaves the metal surface with difficulty, thus slowing down the reaction. It is therefore for the intermediate bond energies that the reaction rate reaches its maximum. The results presented in Figure 5.5 have been obtained in highly purified solutions and, therefore, they highlight the role played by the metal. Under other conditions, the value of i_0 varies greatly depending on the type

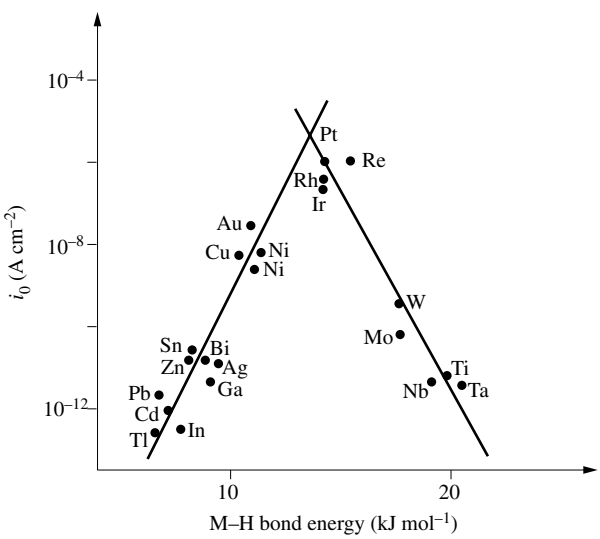


Figure 5.5 Variation of hydrogen electrode exchange current density on different metals as a function of M-H bond energy in 1 M acid solution (adapted from ref. [4]).

© 2007, first edition, EPFL Press

and concentration of impurities, organic or metallic, present in the electrolyte. Prediction under these circumstances is difficult, even impossible.

5.1.4 Determination of reaction order

The ***reaction order*** of a given species that participates in an electrode reaction corresponds to the exponent associated with its concentration in the kinetic law. Let's consider an arbitrary electrode reaction:

$$n\ e\ =\ \Sigma\ \nu_{\mathrm{i}} \cdot B_{\mathrm{i}} \tag{5.54}$$

where ν_{i} represents the stoichiometric coefficient of each species participating in the reaction. Signs follow the convention adopted in Chapter 2.

$$\nu_{\mathrm{i}}\ =\ \nu_{\mathrm{i,red}} \tag{5.55}$$

$$\nu_{\mathrm{i}}\ =\ -\ \nu_{\mathrm{i,ox}} \tag{5.56}$$

The general Butler-Volmer equation for electrode reaction (5.54) is given by the formula:

$$i = i_{\mathrm{a}} + i_{\mathrm{c}} = nFk_{\mathrm{a}} \prod c_{\mathrm{i}}^{p_{\mathrm{a,i}}} \exp\left(\alpha_{\mathrm{a}} nfE\right) - nFk_{\mathrm{c}} \prod c_{\mathrm{i}}^{p_{\mathrm{c,i}}} \exp\left(-\alpha_{\mathrm{c}} nfE\right) \tag{5.57}$$

Here, $p_{\mathrm{a,i}}$ and $p_{\mathrm{c,i}}$ represent the reaction orders of the different species participating in the anodic or cathodic partial reactions. The value of the exponents $p_{\mathrm{a,i}}$ and $p_{\mathrm{c,i}}$ often differ from the stoichiometric coefficients ν_{i}, because the same species can participate in both anodic and cathodic partial reactions.

In order to experimentally determine the reaction order of a given species i, two methods are used. The first is based on equations (5.58) and (5.59), which define, respectively, the anodic and cathodic reaction order of the species. They show that we can obtain $p_{\mathrm{a,i}}$ and $p_{\mathrm{c,i}}$ by measuring the shift of the anodic and cathodic Tafel lines as a function of the concentration of species i, when the concentrations of the other species are maintained constant.

$$p_{\mathrm{a,i}} = \left[\frac{\partial \ln i_{\mathrm{a}}}{\partial \ln c_{\mathrm{i}}}\right]_{\mathrm{E,c \neq c_i}} \tag{5.58}$$

$$p_{\mathrm{c,i}} = \left[-\frac{\partial \ln i_{\mathrm{c}}}{\partial \ln c_{\mathrm{i}}}\right]_{\mathrm{E,c \neq c_i}} \tag{5.59}$$

The second method consists of studying the variation of the exchange-current density i_0 with concentration. The exchange-current density of a given electrode reaction is obtained by setting in (5.57): $i = 0$, $E = E_{\mathrm{rev}}$

$$nFk_{\mathrm{a}} \prod c_{\mathrm{i}}^{p_{\mathrm{a,i}}} \exp\left(\alpha_{\mathrm{a}} nfE_{\mathrm{rev}}\right) = nFk_{\mathrm{c}} \prod c_{\mathrm{i}}^{p_{\mathrm{c,i}}} \exp\left(-\alpha_{\mathrm{c}} nfE_{\mathrm{rev}}\right) = i_0 \tag{5.60}$$

© 2007, first edition, EPFL Press

The following relation then gives the variation of i_0 with the concentration of a species i, when all other concentrations remain constant:

$$\left[\frac{\partial \ln i_0}{\partial \ln c_i}\right]_{c\neq c_i} = p_{a,i} + \alpha_a n f\left[\frac{\partial E_{rev}}{\partial \ln c_i}\right]_{c\neq c_i} = p_{c,i} - \alpha_c n f\left[\frac{\partial E_{rev}}{\partial \ln c_i}\right]_{c\neq c_i} \tag{5.61}$$

By replacing the activities with concentrations in the Nernst equation,

$$E_{rev} = E^0 + \frac{1}{nf}\ln\prod c_i^{-\nu_i} \tag{5.62}$$

we get, by differentiation,

$$\left[\frac{\partial E_{rev}}{\partial \ln c_i}\right]_{c\neq c_i} = -\frac{\nu_i}{nf} \tag{5.63}$$

which yields, with (5.61):

$$\left[\frac{\partial \ln i_0}{\partial \ln c_i}\right]_{c\neq c_i} = p_{a,i} - \alpha_a \nu_i = p_{c,i} + \alpha_c \nu_i \tag{5.64}$$

Example: anodic dissolution of zinc

Zinc is soluble in mercury forming an amalgam. Upon anodic polarization in alkaline solution zinc selectively dissolves from the amalgam [5]:

$$Zn_{(Hg)} + 4\,OH^- \rightarrow Zn(OH)_4^{2-} + 2\,e \tag{5.65}$$

Figure 5.6 shows the exchange-current density i_0 as a function of the zinc concentration in the mercury and of the concentrations of OH^- and $Zn(OH)_4^{2-}$ ions in the electrolyte [5]. These results provide the following values:

$$\left[\frac{\partial \ln i_0}{\partial \ln c_{Zn}}\right]_{c\neq c_{Zn}} \approx 0.5 \tag{5.66}$$

$$\left[\frac{\partial \ln i_0}{\partial \ln c_{OH^-}}\right]_{c\neq c_{OH^-}} = 0 \tag{5.67}$$

$$\left[\frac{\partial \ln i_0}{\partial \ln c_{Zn(OH)_4^{2-}}}\right]_{c\neq c_{Zn(OH)_4^{2-}}} \approx 0.5 \tag{5.68}$$

© 2007, first edition, EPFL Press

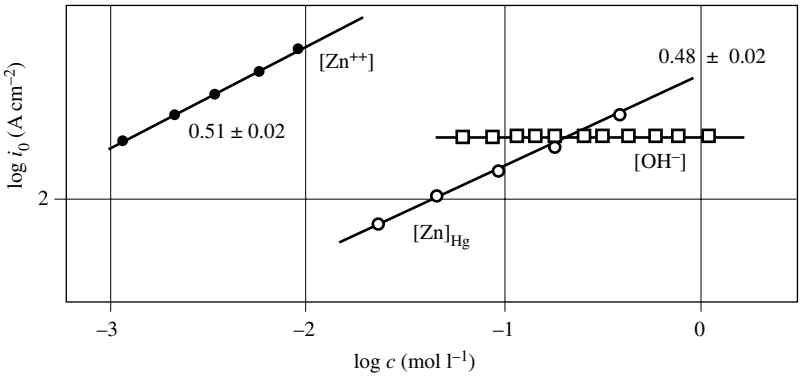


Figure 5.6 Variation of the exchange-current density of the Zn(Hg)/$Zn(OH)_4^{2-}$ electrode as a function of the concentrations of the different reactants and products [5].

The relation (5.64), where we assign the value of 0.5 to α_a, allows us to calculate the reaction order for the species Zn, $Zn(OH)_4^{2-}$ and OH^- for the anodic reaction.

$$\text{Zn:} \quad 0.5 = p_{a,Zn} - 0.5\,(1) \rightarrow p_{a,Zn} = 1$$

$$OH^-: \quad 0 = p_{a,OH^-} - 0.5\,(4) \rightarrow p_{a,OH^-} = 2$$

$$Zn^{2+}: \quad 0.5 = p_{a,Zn(OH)_4^{2-}} - 0.5\,(-1) \rightarrow p_{a,Zn(OH)_4^{2-}} = 0$$

These values suggest the following reaction mechanism: in a first step that is rate determining, the weakly soluble species $Zn(OH)_2$ is formed, followed by a dissolution and complexation reaction that produces the anion $Zn(OH)_4^{2-}$.

$$Zn_{(Hg)} + 2\,OH^- \rightarrow Zn(OH)_2 + 2\,e \quad \text{(RDS)} \tag{5.69}$$

$$Zn(OH)_2 + 2\,OH^- \rightarrow Zn(OH)_4^{2-} \tag{5.70}$$

5.2 ELECTROCHEMICAL NON-STEADY-STATE METHODS

5.2.1 Overview of electrochemical non-steady-state methods

Non-steady-state electrochemical techniques are widely used in corrosion science and engineering for the study of corrosion and protection mechanisms and to measure corrosion rate. In contrast to steady state polarization methods, transient techniques are able to distinguish reaction phenomena with different time constants. For example,

© 2007, first edition, EPFL Press

the relaxation time of charge transfer reactions at the electrode-electrolyte interface is often smaller than that of transport phenomena in the electrolyte.

Table 5.7 gives an overview of the most commonly used electrochemical non-steady-state methods (in parentheses: alternative names). Typically, one imposes a variation in either the current or the potential at the working electrode and then records its response as a function of time or frequency.

Table 5.7 Electrochemical non steady state methods.

Imposed quantity	Form of applied signal	Measured quantity	Method
E	step	$I = f(t)$	potential step method (potentiostatic transient method)
I	step	$E = f(t)$	current step method (galvanostatic transient method)
E, dE/dt	linear sweep	$I = f(E, \mathrm{d}E/\mathrm{dt})$	potential sweep method (linear sweep voltammetry)
E, dE/dt	cyclic sweep	$I = f(E, \mathrm{d}E/\mathrm{dt})$	cyclic voltammetry
$\Delta E = \lvert\Delta E\rvert \sin(\omega t)$	sine wave	$Z = f(\omega)$	impedance method (impedance spectroscopy)

5.2.2 Potential step method

A potential step is applied to the electrode, and the variation in current is measured as a function of time. This method has proven particularly useful for the identification of transport processes in the electrolyte and for studying the growth of passive oxide films (Chapter 6).

Here, we develop the principle of the potential step method for a particularly simple reaction, namely, the electrodeposition of a metal under conditions where the deposition rate is controlled by mass transport.

$$\mathrm{M}^{n+} + n\,e \rightarrow \mathrm{M} \tag{5.71}$$

Assuming that the charge transfer reaction at the interface is sufficiently rapid for a quasi-equilibrium state to exist, the electrode potential follows the Nernst equation. If, in addition, the activity of the ion M^{n+} is equal to its concentration the following relation holds for the potential:

$$E = E^0 + \frac{RT}{nF} \ln c_s \tag{5.72}$$

Here c_s indicates the concentration at the metal surface. At equilibrium, the value of c_s is equal to that of the bulk concentration c_b, but when a negative potential step is applied, copper-ions react at the metal-solution interface and as a consequence, a

© 2007, first edition, EPFL Press

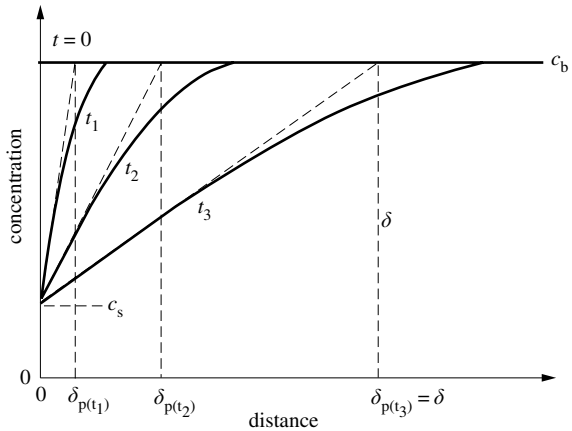


Figure 5.8 Concentration profiles near an electrode after the application of a potential step. The steady state is reached when the thickness of the non-steady-state diffusion layer, $\delta_{p(t)}$, becomes equal to δ, the thickness of the steady-state diffusion layer.

concentration gradient is established in the vicinity of the electrode (Figure 5.8). Assuming for simplicity that there are no migration effects, the cathodic current density i_c is proportional to the concentration gradient at the electrode surface (Section 4.3).

$$i_c = -nFD \left.\frac{dc}{dy}\right|_{y=0} \tag{5.73}$$

To know i_c, we therefore must calculate the evolution of the copper ion concentration near the electrode as a function of time and distance. For this, we use ***Fick's second law***:

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial y^2} \tag{5.74}$$

with the initial conditions:

$$t = 0: \quad c = c_b \tag{5.75}$$

and the boundary conditions:

$$t \geq 0\,,\ y = \infty : \ c = c_b \tag{5.76}$$

$$t \geq 0\,,\ y = 0 : \ c = c_s \tag{5.77}$$

© 2007, first edition, EPFL Press

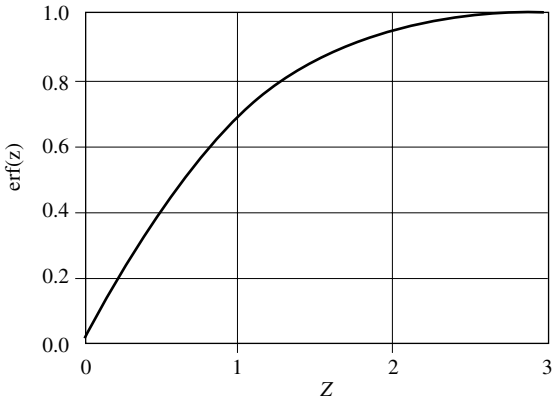


Figure 5.9 The error function erf(z).

The integration of (5.74) yields

$$\frac{c - c_{\mathrm{b}}}{c_{\mathrm{s}} - c_{\mathrm{b}}} = 1 - \mathrm{erf}(z) \tag{5.78}$$

where the error function erf*(z)* is defined as

$$\mathrm{erf}(z) = \frac{2}{\pi^{1/2}} \int_0^z \exp(-z'^2)\,\mathrm{d}z' \tag{5.79}$$

and the dimensionless variable z corresponds to the expression

$$z = \frac{y}{(4Dt)^{1/2}} \tag{5.80}$$

Figure 5.9 gives the value of erf*(z)* as a function of z. The derivative of (5.78) at the electrode surface provides, along with (5.73), a relation between current and time, known as ***Cottrell equation***.

$$I(t) = i(t)A = -nFA(c_b - c_s)\left(\frac{D}{\pi t}\right)^{1/2} \tag{5.81}$$

This equation describes the variation of the current following a potential step for an electrochemical reaction controlled by diffusion. It predicts that the current decreases proportionately to $1/t^{1/2}$. By introducing a ***non-steady-state diffusion layer***,

$$\delta_{\mathrm{p(t)}} = (\pi D t)^{1/2} \tag{5.82}$$

© 2007, first edition, EPFL Press

we obtain

$$I(t) = -nFAD\,\frac{c_b - c_s}{\delta_{p(t)}} \tag{5.83}$$

This equation highlights the progressive growth of the diffusion layer in the absence of convection (Figure 5.8). For a reaction controlled only by diffusion, the current never reaches steady state when a potential step is applied.

In convective systems, the situation is different, because the thickness of the non-steady-state diffusion layer, $\delta_{p(t)}$, is limited by that of the steady-state diffusion layer δ, which depends only on the prevailing hydrodynamic conditions (see Sect. 4.3). As the non-steady-state diffusion layer grows, convection contributes increasingly to the transport of the reacting species. Eventually, when $\delta_{p(t)}$ becomes comparable to the steady-state diffusion layer thickness δ the concentration profile no longer changes with time.

Equation (5.84) defines the dimensionless parameter τ^*_{pot} which characterizes, in forced convection systems, the transition from the non-steady-state to the steady state transport regimes.

$$\tau^*_{pot} = \frac{\delta_{p(t)}}{\delta} = \frac{(\pi D t)^{1/2}}{\delta} \tag{5.84}$$

In the non-steady-state regime: $\tau^*_{pot} << 1$, in the steady-state regime: $\tau^*_{pot} >> 1$. By setting $t = \tau_p$ for $\delta_{p(t)}/\delta = 1$, we can define the relaxation time τ_p of the concentration profile resulting from application of a potential step in a forced convection system:

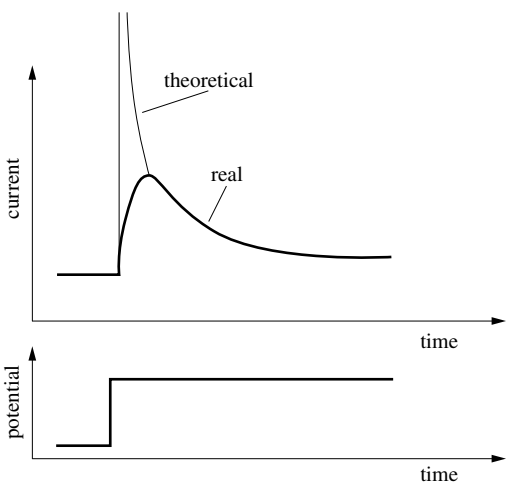


Figure 5.10 Current transient following application of a potential step.

© 2007, first edition, EPFL Press

$$\tau_p = \frac{\delta^2}{\pi D} \tag{5.85}$$

This result shows that the steady state is more quickly established when convection is strong, in other words, when δ is small.

The Cottrell equation predicts an infinite current for $t = 0$, immediately following the potential jump. In practice, several factors limit the current and prevent such a situation from ever occurring: the limited rise time and current output of the potentiostat, the ohmic drop in the solution that prevents the potential to instantly reach the prefixed value, and the kinetic limitations of the interfacial reaction. In potentiostatic step experiments, the current therefore attains the theoretical value given by the Cottrell equation only after a certain time lag. Figure 5.10 shows schematically the typical shape of the current transients observed in potential step experiments.

5.2.3 Current step method

In the current step or *galvanostatic transient method*, one applies a current step to the working electrode and then records the potential as a function of time. Three phenomena contribute to the shape of the observed potential transients:

- the mass transport of the reacting species;
- the charging of the double layer;
- the potential drop in the electrolyte.

Mass transport

In order to calculate how mass transport affects the potential transient, one has to solve Fick's equation (5.74) and replace the boundary condition (5.77) with (5.86).

$$t \geq 0, \quad y = 0; \quad \frac{\partial c}{\partial y} = -\frac{I}{nFAD} = constant \tag{5.86}$$

Figure 5.11 shows the evolution of the concentration profile of a reactant near the electrode, following the application of a current step. The concentration gradient at the electrode surface remains constant, while the surface concentration, given by equation (5.87), decreases.

$$c_s - c_b = \frac{2It^{1/2}}{nFA(\pi D)^{1/2}} \tag{5.87}$$

Eventually, the surface concentration of the reacting species, c_s, reaches zero. The corresponding time, τ_g, called the ***transition time***, is easily evaluated from (5.87):

$$\tau_g = \frac{n^2F^2A^2c_b^2\pi D}{4I^2} \tag{5.88}$$

© 2007, first edition, EPFL Press

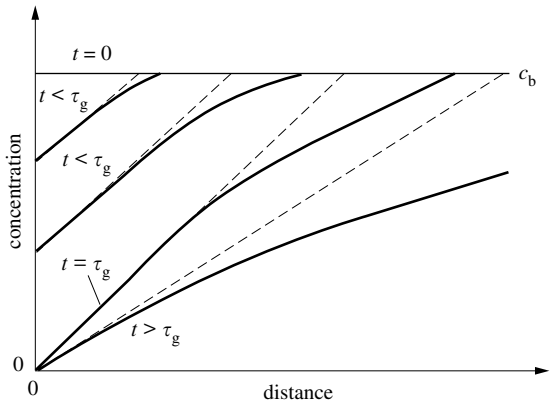


Figure 5.11 Concentration profile of a reactant near the electrode surface upon application of a current step. At the transition time $t = \tau_g$, the concentration at the electrode surface drops to zero.

Equation (5.88) is known as the ***Sand equation.*** Beyond the transition time, $t > \tau_g$, the concentration gradient at the surface progressively decreases. Under these conditions a second reaction increasingly accounts for the applied current (see below).

The decrease in surface concentration of the reacting species creates an overpotential (Sect. 4.3):

$$\eta = \frac{RT}{nF} \ln \frac{c_s}{c_b} \tag{5.89}$$

Division of (5.87) by c_b gives:

$$\frac{c_s}{c_b} = 1 + \frac{2\,I t^{1/2}}{n\,F\,A\,c_b\left(\pi D\right)^{1/2}} \tag{5.90}$$

By plugging this equation into (5.89), the variation of the overpotential as a function of time is obtained:

$$\eta = \frac{RT}{nF} \ln\left[1 + \frac{2\,I t^{1/2}}{n\,F\,A\,c_b\left(\pi D\right)^{1/2}}\right] \tag{5.91}$$

For a cathodic reaction ($I < 0$), this equation predicts an infinitely negative overpotential at the transition time, when the reactant concentration at the surface becomes zero. In reality, one observes a sudden increase of the overpotential at the transition time, but its value never reaches infinity. The reason is that other reactions such as water decomposition become possible at higher overvoltage and they then determine the electrode potential. Figure 5.12 schematically shows a galvanostatic potential transient that exhibits a potential rise at the transition time. The measurement of the transition

© 2007, first edition, EPFL Press

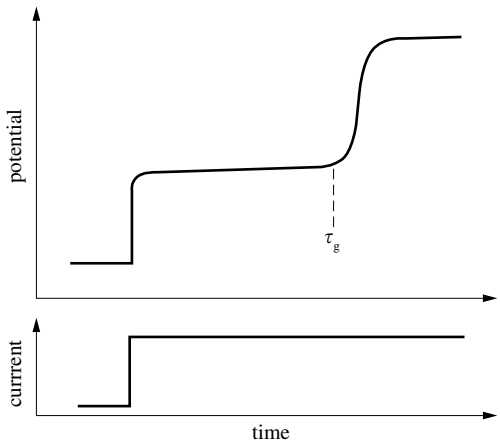


Figure 5.12 Potential transient following application of a current step.

time allows one to obtain a number of parameters, in particular the diffusion coefficient of the reactant, provided its concentration in the bulk electrolyte is known.

The Sand equation is applicable to cathodic deposition as well as to anodic dissolution reactions. During dissolution, the concentration of the ions produced at the surface increases with time. When it reaches a critical concentration at the time $t = \tau_g$, reaction products may precipitate at the surface and form a film. (The same behavior was discussed in subsection 4.3.4 for steady-state conditions). By analogy to (5.88) the Sand equation of an anodic dissolution reaction is:

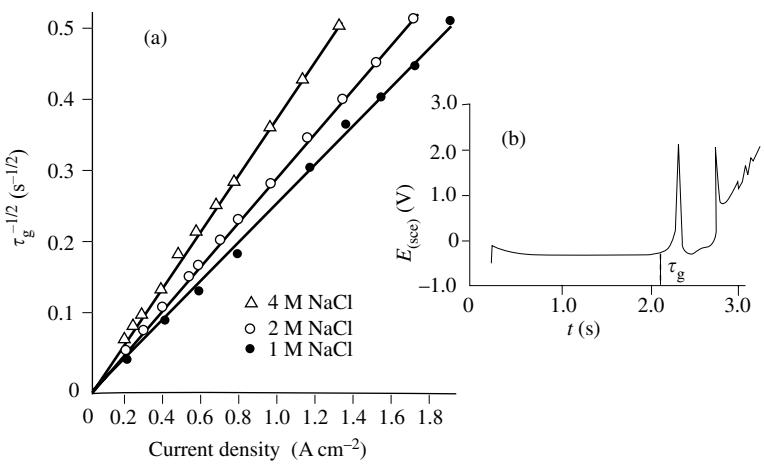


Figure 5.13 (a) Variation of transition time t_g as a function of current density for anodic dissolution of Fe in 4M NaCl, in the absence of convection ; (b) View of a potential transient.

© 2007, first edition, EPFL Press

$$\tau_g = \frac{n^2 F^2 A^2 \left(c_{\text{crit}} - c_{\text{b}}\right)^2 \pi D}{4 I^2} \tag{5.92}$$

Under steady-state conditions, the critical concentration c_{crit} corresponds to the saturation concentration c_{sat}. Under non-steady-state conditions, such as those prevailing in current step experiments, the precipitation of a reaction product film most often requires a certain degree of supersaturation and therefore the value of the concentration c_{crit} can exceed the saturation concentration c_{sat}.

Figure 5.13 presents a gavanostatic potential transient, measured during the dissolution of an iron electrode immersed in a concentrated solution of sodium chloride. At the transition time, the Fe^{2+} concentration at the anode surface reaches the critical value for precipitation of an $FeCl_2$ film. The phenomenon leads to a sharp potential rise followed by potential fluctuations related to the behavior of the salt-film formed. The observed relation $i \propto (1/\tau_g)^{1/2}$ is in agreement with equation (5.92).

Double layer capacity

The gavanostatic transient method can be used to measure the double layer capacity and the ohmic drop in the electrolyte between the working and reference electrodes. Figure 5.14 shows the initial variation of the potential resulting from a current step. During this period, diffusion is of little importance because its time constant is usually much larger. The electrode acts like a capacitor connected in parallel with a resistance, where the capacitor represents the double layer capacity and the parallel resistance corresponds to the transfer resistance of the electrode reaction (Section 3.5). The applied current I is the sum of the capacitive current I_C given by (5.94) and the faradaic current I_F described by the Butler-Volmer equation:

$$I = I_C + I_F \tag{5.93}$$

$$I_C = C \frac{dE}{dt} \tag{5.94}$$

When the current is turned on, the faradaic current is zero, because the double layer must be charged first.

$$t = 0 \ : \qquad I = I_C \tag{5.95}$$

As the potential across the double layer increases, the faradaic current also increases, while the capacitive current becomes less important. The steady state is reached when the capacitive current becomes zero.

According to (5.94) and (5.95) the derivative of the potential with respect to time at $t = 0$ provides the double layer capacity. The initial potential jump represents the quantity $\Delta\Phi_\Omega = IR_\Omega$, because the response of an ohmic element is in principle instantaneous. In this way we can find the ohmic resistance, R_Ω, in the electrolyte between the Luggin capillary and the working electrode.

© 2007, first edition, EPFL Press

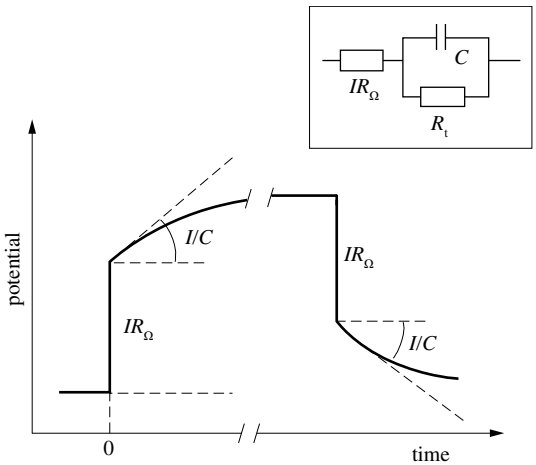


Figure 5.14 Measurement of double layer capacity by galvanostatic transient method: switch-on and switch-off of current.

Alternatively, we can also determine $\Delta\Phi_\Omega$ and C from the potential decay following current switch-off at the time $t = t'$. At open circuit, the double layer discharges across the resistance R_t while a faradaic reaction takes place.

$$I_f = -C\,\frac{\mathrm{d}E}{\mathrm{d}t} \tag{5.96}$$

At $t = t'$ the faradaic current is equal to the applied current $I_f = I$. The derivative of the potential with respect to time, at current switch-off, therefore permits to evaluate the double layer capacity. Figure 5.14 shows an idealized potential transient. Unfortunately, in practice, other capacitances present in the electrochemical system often attenuate the sharpness of the potential jump and thus complicate the distinction between ohmic and capacitive contributions.

5.2.4 Voltammetry

In voltammetry one sweeps the potential at a constant rate across a predetermined interval. The technique is therefore also referred to as ***linear sweep voltammetry***.

$$E(t) = E(t=0) + \lambda\, t \tag{5.97}$$

$$\lambda = \frac{\mathrm{d}E}{\mathrm{d}t} = constant$$

The measured current is a function of the sweep rate λ and the time: $I = f(t, \lambda)$, but often it is plotted as a function of potential $I = f(E, \lambda)$. The experimental curve $I = f(E, \lambda)$ is called a ***voltammogram*** or *linear sweep voltammogram*.

© 2007, first edition, EPFL Press

In voltammetry, a distinction is made between two operating modes:

- single sweep voltammetry, which consists of a single potential sweep;
- cyclic voltammetry, which consists of one or several bidirectional sweeps.

Single sweep voltammetry

Let us first look at the influence of mass transport on the function $I = f(E, \lambda)$ in a forced convection system. Figure 5.15 schematically shows the current for three different potential sweep rates. We assume that a reactant is consumed at the electrode and that the variation of the potential is due to concentration overpotential only.

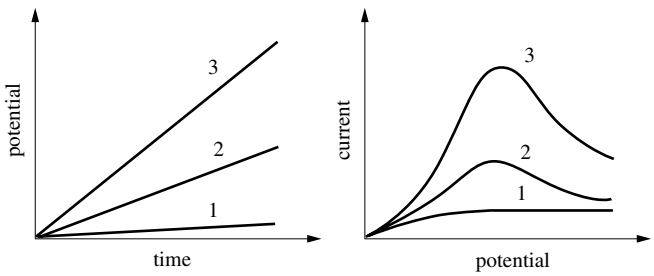


Figure 5.15 Effect of sweep rate on measured current in linear sweep voltammetry.

Curve 1 represents a sweep rate sufficiently slow to permit steady state mass transport conditions to be established at the electrode surface. The measured current therefore corresponds to that of a stationary polarization curve, $I = f(E)$, which exhibits a limiting current plateau. At higher sweep rates, the concentration profile can no longer attain steady state and as a consequence one observes a current maximum (curve 2). The maximum becomes more pronounced as the sweep rate further increases (curve 3). The concentration profiles presented in Figure 5.16 provide an explanation for this behavior. At first, the concentration gradient at the surface increases with time while the surface concentration of the reactant decreases. Under these conditions the concentration gradient at the surface, and hence the current, exceeds the value corresponding to steady state mass transport. Once the surface concentration of the reactant reaches zero, the concentration gradient at the surface progressively decreases and approaches the steady-state value.

A quantitative analysis of this behavior allows us to define a dimensionless time, τ_v^*, which characterizes the transition from non-steady-state to steady-state conditions [7]:

$$\tau_v^* = \frac{D\,R\,T}{\delta^2 n\,F\,\lambda} \tag{5.98}$$

Based on the value of τ_v^*, we can identify two limiting cases:

- $\tau_v^* >> 1$, steady-state mass transport regime: the concentration profile in the diffusion layer corresponds to that of the steady state and the limiting current density depends exclusively on convection conditions.

© 2007, first edition, EPFL Press

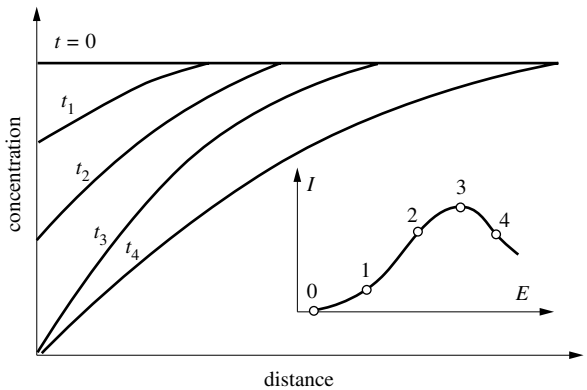


Figure 5.16 Non-steady-state concentration profiles near the electrode during a linear potential sweep.

• $\tau_v^* << 1$, non-steady-state mass transport regime: the concentration profile inside the diffusion layer changes with time and the effect of convection on $I = f(E,\lambda)$ is negligible.

Steady-state mass-transport regime

In corrosion and in electrochemistry, the potential sweep technique is commonly used to measure polarization curves, and the result is referred to as ***potentiodynamic polarization curves***. Equation (5.98) offers a criterion for the selection of the maximum sweep rate while still working under steady state conditions with respect to mass transport. As a rule of thumb, for a value of $\tau_v^* > 20$ the error in the measured steady-state limiting current is less than 1% [7]. Equation (5.98) shows that to attain a steady state, the sweep rate must be the slower the larger the diffusion layer thickness, in other words, the weaker the convection. If, for example, $D = 10^{-9}$ m^2 s^{-1} , $n = 2$ and $\delta = 10$ μm, the sweep rate must not exceed 15 mV s^{-1}.

In corrosion, phenomena other than mass transport in the electrolyte can slow down the establishment of steady state conditions, including adsorption, precipitation or film growth. Especially, solid state transport processes in passive oxide films are generally slow (Chap. 6) and as a consequence the measured current density will depend on the sweep rate, even if from a solution mass transport point of view steady state prevails ($\tau_v^* >>1$). Polarization curves measured under these conditions are sometimes called ***pseudo-steady-state polarization*** curves. When reporting such data one should always indicate the sweep rate used.

Non-steady-state mass transport regime

At high sweep rates ($\tau^*_v << 1$), the contribution of convection to transport is negligible compared to that of non-steady-state diffusion. We can then calculate the

© 2007, first edition, EPFL Press

theoretical shape of the current transient by applying Fick's second law (5.74) with the appropriate boundary conditions. For a reversible reaction, in other words, in the absence of charge transfer overvoltage, the maximum current I_p is given by the ***Randles-Sevcik equation***:

$$I_p = k_{RS}\, nFA\left(\frac{nF}{RT}\right)^{1/2} c_b D^{1/2} \lambda^{1/2} \tag{5.99}$$

where c_b is expressed in mol cm^{-3}, λ in V s^{-1}, I_p in A and the surface A in cm^2. The constant k_{RS}, has the value k_{RS} = 0.4463. The Randles-Sevcik equation is also applicable to reactions that involve a slow rate of charge transfer (irreversible reactions), but the value of k_{RS} in this case depends on the kinetics of the charge transfer reaction. In practice, the proportionality between I_p and the concentration c_b can be used to determine the concentration of an electrochemical reactant if its diffusion constant is known or vice versa to determine D for a reactant of given concentration c_b.

Cyclic voltammetry

In cyclic voltammetry, a potential range is swept in forward and reverse direction once or several times. Figure 5.17 schematically shows the cyclic voltammogram of an electrode reaction involving dissolved species such as for example Fe^{2+} and Fe^{3+} ions. In the anodic sweep Fe^{2+} is oxidized to Fe^{3+}, in the reverse direction the Fe^{3+} is reduced yielding a cathodic current. In simple cases, the shape of the voltammogram and the difference between the potential maxima allow one to identify the reaction mechanism [8], but in corrosion, these criteria are usually of limited usefulness and shall therefore not be further discussed here.

Multiple-sweep cyclic voltammetry is a useful method for the qualitative investigation of complex reaction phenomena at electrodes. In this method, the potential is swept repetitively in the forward and reverse directions at a relatively high rate, typically from one to several volts per second. A high scan rate permits to distinguish phenomena with fast time constants. Furthermore, it improves the

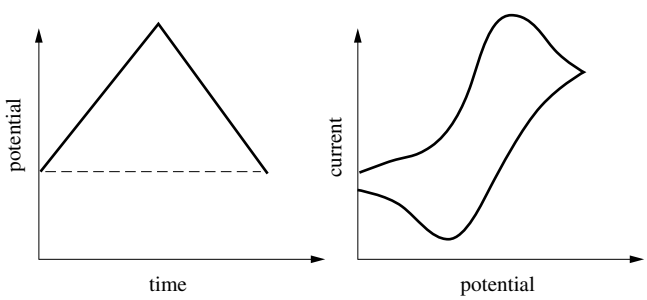


Figure 5.17 Cyclic voltammogram.

© 2007, first edition, EPFL Press

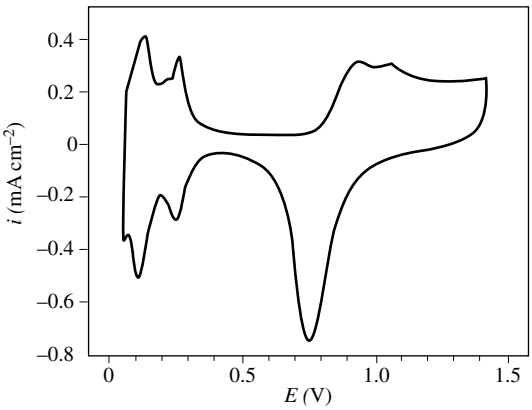


Figure 5.18 Cyclic voltammogram of Pt electrode in deaerated 0.5 M H_2SO_4. Sweep rate: 0.1 V/s [9].

reproducibility of the results by reducing irreproducibility caused by contaminants, because each oxidation-reduction cycle "cleans" the electrode surface.

Multiple-sweep cyclic voltammetry is most powerful for the identification of adsorbed species and intermediate products. Figure 5.18 shows a cyclic voltammogram obtained with a platinum electrode immersed in 0.5 M sulfuric acid.
By following the current in the direction of increasing potential, we observe two anodic peaks, which represent the oxidation of adsorbed hydrogen:

$$H_{ads} \rightarrow H^+ + e \tag{5.100}$$

The presence of two peaks indicates that the adsorbed hydrogen assumes two energy states, depending on the adsorption site. At higher potential, between 0.3 and 0.8 V, one observes a plateau of relatively low current. Apparently, no faradaic reaction takes place in this potential range and all of the current serves to charge the electric double layer. The area under the curve in the region of the plateau therefore permits the evaluation of the double layer charge Q_{dl}.

$$\int I\,dt = \int I\left(\frac{dt}{dE}\right)dE = \frac{1}{\lambda}\int I\,dE = Q_{dl} \tag{5.101}$$

Differentiation with respect to the potential yields the double layer capacity C:

$$C = \frac{dQ_{dl}}{dE} = \frac{I}{\lambda} \tag{5.102}$$

In Figure 5.18, the current plateau (I = constant) reveals that the capacity remains the same throughout this part of the diagram.

Between 0.8 V and 1.2 V the anodic current increases as a result of the oxidation of platinum to platinum oxide. When the sweep direction is reversed, the reduction of the oxide causes a cathodic peak at 0.75 V. At more negative potentials, a current

© 2007, first edition, EPFL Press

plateau appears that corresponds to charging the double layer with reversed polarity. The two peaks that follow represent the adsorption of hydrogen. By integrating the current over the appropriate potential range we can determine the electric charge that is associated with each of these phenomena.

A steady-state polarization curve would show no current at all in the potential range between hydrogen and oxygen formation. The described transient phenomena taking place on platinum in the potential region shown can only be observed by a non-steady state technique. The results of Figure 5.18 suggest that proton reduction involves different steps and that adsorbed atomic hydrogen is indeed an intermediate reaction product (cf. Section 5.1).

5.2.5 Impedance Method

The impedance method consists in measuring the response of an electrode to a sinusoidal potential modulation of small amplitude (typically 5-10 mV) at different frequencies. The *ac* modulation is superimposed either onto an applied anodic or cathodic potential or onto the corrosion potential. Another possibility is to modulate the current and to measure the potential. Impedance measurements as a function of modulation frequency are commonly referred to in the literature as ***electrochemical impedance spectroscopy***, abbreviated ***EIS***. Among the different transient methods discussed in this chapter, EIS is most widely used in corrosion studies. It serves for the measurement of uniform corrosion rates, for the elucidation of reaction mechanisms, for the characterization of surface films and for testing of coatings.

Electrode response to a sinusoidal perturbation of the potential

A sinusoidal modulation of weak amplitude $\Delta E = E - E_{st}$ is superimposed onto the steady-state potential E_{st} of an electrode.

$$\Delta E = |\Delta E| \sin(\omega t) \tag{5.103}$$

Here, $|\Delta E|$ is the amplitude, $\omega = 2\pi f$ is the radial frequency in radians per second (f is the frequency expressed in hertz). In order to maintain a linear response of the electrode, the modulation amplitude must not exceed about 10 mV. The sinusoidal perturbation of the potential induces a sinusoidal current ΔI, superimposed onto the steady state current, with the phase shift of ϕ with respect to the potential.

$$\Delta I = |\Delta I| \sin(\omega t') = |\Delta I| \sin(\omega t - \phi) \tag{5.104}$$

$$\phi = \omega (t - t')$$

Figure 5.19 shows the normalized ΔE and ΔI functions. In polar coordinates, these functions are represented by vectors of length $|\Delta E|$ and $|\Delta I|$ rotating counterclockwise at the radial frequency ω. These functions are most easily described by using of complex numbers.

© 2007, first edition, EPFL Press

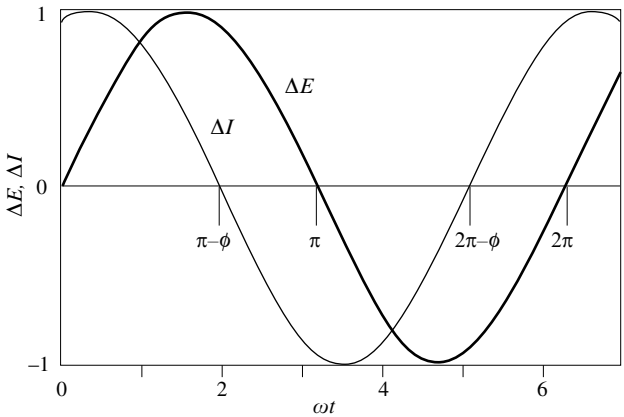


Figure 5.19 Phase lag of ac current with respect to potential.

In the complex plane (Figure 5.20), a sinusoidal perturbation to the potential is represented by a vector ΔE , which corresponds to the sum of a real term and an imaginary term.

$$\Delta E = \Delta E_{\mathrm{Re}} + j\,\Delta E_{\mathrm{Im}} \tag{5.105}$$

Here, $j = (-1)^{1/2}$, ΔE_{Re} and ΔE_{Im} designate, respectively, the real and imaginary parts of ΔE . The use of Figure 5.20 allows us to derive the following relations:

$$\Delta E_{\mathrm{Re}} = |\Delta E| \cos \omega t \tag{5.106}$$

$$\Delta E_{\mathrm{Im}} = |\Delta E| \sin \omega t \tag{5.107}$$

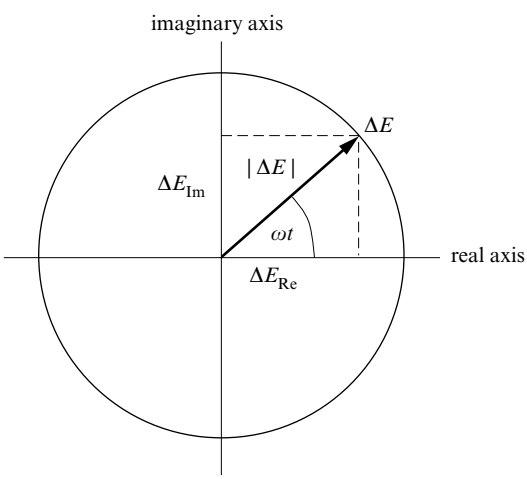


Figure 5.20 Representation of the vector ΔE in the complex plane.

© 2007, first edition, EPFL Press

With the mathematical identity,

$$\cos \omega t + j \sin \omega t = \exp(j\omega t) \tag{5.108}$$

we obtain

$$\Delta E = |\Delta E| \exp(jwt) \tag{5.109}$$

In an analogous fashion:

$$\Delta I = |\Delta I| \exp(j(\omega t - \phi)) \tag{5.110}$$

The following equation defines the ***electrochemical impedance*** Z :

$$Z = \frac{\Delta E}{\Delta I} = \frac{|\Delta E| \exp(j\,\omega t)}{|\Delta I| \exp(j\,(\omega t - \phi))} \tag{5.111}$$

By setting

$$|Z| = \frac{|\Delta E|}{|\Delta I|} \tag{5.112}$$

we obtain

$$Z = |Z| \exp(j\phi) \tag{5.113}$$

In the complex plane, the impedance Z is represented by a vector, characterized by the ***modulus*** $|\Delta Z|$ and the ***phase shift*** ϕ (Figure 5.21). It is also possible to represent Z as the vector sum of its real and imaginary parts:

$$Z = Z_{\mathrm{Re}} + j\, Z_{\mathrm{Im}} \tag{5.114}$$

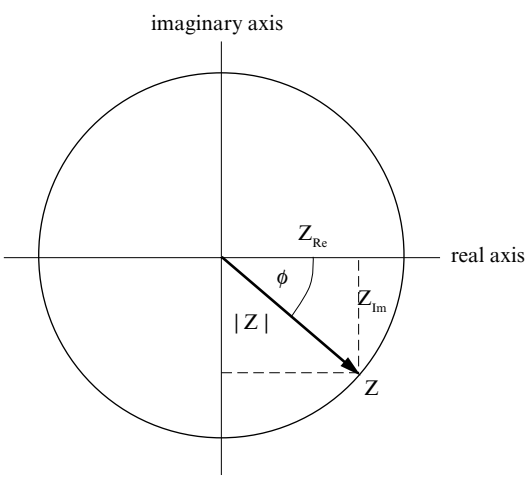


Figure 5.21 Vector representation of impedance in the complex plane.

© 2007, first edition, EPFL Press

The modulus can be calculated using Pythagoras' theorem:

$$|Z| = (Z_{\mathrm{Re}}^2 + Z_{\mathrm{Im}}^2)^{1/2} \tag{5.115}$$

According to Figure 5.21:

$$\tan\phi = \frac{Z_{\mathrm{Im}}}{Z_{\mathrm{Re}}} \tag{5.116}$$

The phase shift is therefore equal to:

$$\phi = \arctan\left(Z_{\mathrm{Im}} / Z_{\mathrm{Re}}\right) \tag{5.117}$$

Experimental determination of the electrochemical impedance

Several techniques are available for the measurement of the impedance of an electrical circuit [10]. Most often today, a ***transfer function analyzer*** *(TFA)* is employed.

Figure 5.22 shows the principle of the method. A generator supplies a sinusoidal signal $x(t) = x_0 \sin \omega t$. The system responds with a signal $S(t)$ that differs from $x(t)$ by its phase and amplitude. The analyzer multiplies $S(t)$ with a reference signal, either in phase with $x(t)$ or shifted by 90°. Integration between 0 and t', t' being a multiple of the signal period, yields the real part S_{Re} and the imaginary part S_{Im} of the signal $S(t)$ [10].

$$S_{\mathrm{Re}} = \frac{1}{t'} \int_0^{t'} S(t) \sin(\omega t)\, \mathrm{d}t \tag{5.118}$$

$$S_{\mathrm{tm}} = \frac{1}{t'} \int_0^{t'} S(t) \cos(\omega t)\, \mathrm{d}t \tag{5.119}$$

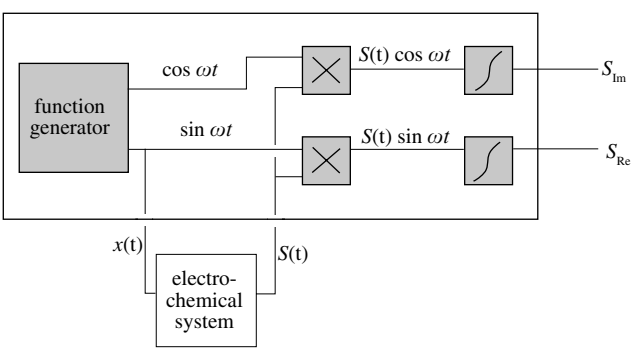


Figure 5.22 Transfer function analyzer (schematic of the operation).

© 2007, first edition, EPFL Press

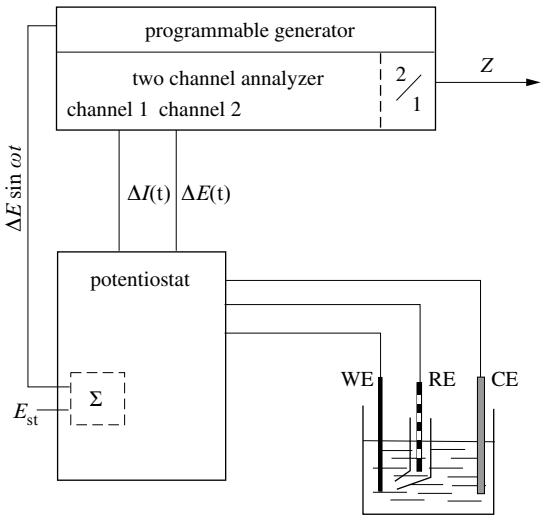


Figure 5.23 Impedance measurement by use of a transfer function analyser under potentiostatic control.

The integration eliminates the harmonics of $S(t)$ and the background noise, provided the integration time t' is sufficiently long. In practice, a compromise is made between the desired precision of the results and the acceptable length of integration time.

In a typical impedance experiment, a potentiostat supplies the steady-state potential to which a sinusoidal perturbation is superposed by a programmable frequency generator. The latter is often built into a two-channel transfer function analyzer (Figure 5.23), thus permitting simultaneous measurement of the potential and the current. The analyzer determines the real and imaginary parts of the two quantities and by division calculates the impedance Z of the electrochemical system. Using appropriate software the data are transferred into a computer memory and stored for subsequent drawing of impedance diagrams. The frequency range employed in impedance measurements typically goes from several millihertz to a hundred kilohertz or more. At low frequency, the experiments require much time and there is the risk that the electrode surface might change. The maximum usable frequency is determined by the time response of the potentiostat and the capacitances associated with the cell and the electric circuit.

Impedance of the passive elements

Before discussing the impedance of electrochemical systems it is useful to recall briefly the alternating current response of electrical circuit elements. Three passive elements are normally present in an electrical circuit:

- the resistance R;
- the capacitance C;
- the inductance L.

© 2007, first edition, EPFL Press

A sinusoidal voltage $\Delta E = |\Delta E| \sin(\omega t)$ applied to an ohmic resistance induces an alternating current that is in phase with the applied voltage:

$$\Delta I = |\Delta I| \sin(\omega t),$$

The impedance of an ohmic circuit element is therefore equal to its resistance R:

$$Z = \frac{\Delta E}{\Delta I} = R \tag{5.120}$$

The impedance of a capacitor can be calculated with the relation:

$$\Delta I = C \frac{\mathrm{d}\Delta E}{\mathrm{d}t} \tag{5.121}$$

For the sinusoidal voltage function (5.103) and with the identity $\cos \omega t = \sin(\omega t + \pi/2)$ we obtain:

$$\Delta I = C\omega |\Delta E| \cos(\omega t) = C\omega |\Delta E| \sin(\omega t + \pi/2) \tag{5.122}$$

By comparing this to the relation (5.104), we find for the phase shift and the modulus:

$$\phi = -\pi/2 \tag{5.123}$$

$$|\Delta Z| = \frac{|\Delta E|}{|\Delta I|} = \frac{1}{\omega C} \tag{5.124}$$

The current through a capacitor exhibits a positive phase shift of 90° relative to the voltage. It is often more practical to express the impedance of a capacitor, Z_C, in complex form. For this we write:

$$Z_C = |Z| \exp(j\phi) = (1/\omega C) \exp(-j\pi/2)$$

$$= (1/\omega C) [\cos(\pi/2) - j \sin(\pi/2)] \tag{5.125}$$

Since $\cos(\pi/2) = 0$ and $\sin(\pi/2) = 1$ we find:

$$Z_C = -\frac{j}{\omega C} \tag{5.126}$$

In an analogous fashion for the inductance, we can write:

$$\Delta E = L \frac{\mathrm{d}\Delta I}{\mathrm{d}t} \tag{5.127}$$

$$\Delta I = \frac{1}{L} \int \Delta E \, \mathrm{d}t = -\frac{|\Delta E|}{\omega L} \cos(\omega t) \tag{5.128}$$

© 2007, first edition, EPFL Press

with the identity cos (ωt) = – sin ($\omega t - \pi/2$):

$$\Delta I = \frac{|\Delta E|}{\omega L} \sin\left(\omega t - \pi/2\right) \tag{5.129}$$

By setting (5.129) and (5.104) equal, we obtain the phase shift and the modulus:

$$\phi = \pi/2$$

$$|\Delta Z| = \omega L.$$

In complex form, the impedance of an inductance is written as:

$$\begin{aligned} Z_I &= |Z| \exp(j\phi) \\ &= (\omega L) \exp(j\pi/2) = (\omega L) [\cos(\pi/2) + j \sin(\pi/2)] \end{aligned} \tag{5.130}$$

and therefore

$$Z_I = j\omega L \tag{5.131}$$

In Table 5.24, the impedances and admittances of the three passive elements are given. The ***admittance*** A_ω is the inverse of the impedance.

$$A_\omega = \frac{1}{Z} \tag{5.132}$$

Table 5.24 Impedance and admittance of passive elements.

Element	Z_{Re}	Z_{Im}	A_{Re}	A_{Im}
Resistance	R	0	$1/R$	0
Capacitance	0	$-j/\omega C$	0	$-j\omega C$
Inductance	0	$j\omega L$	0	$j/\omega L$

Impedance of an electrical circuit

For an alternating current, the electrical impedance plays a similar role as the ohmic resistance for a direct current. The impedance of an electrical circuit composed of passive elements connected in series is equal to sum of the impedances of the individual elements, Z_k,

$$Z = \Sigma Z_k \tag{5.133}$$

and for a circuit with elements connected in parallel:

$$\frac{1}{Z} = \sum \frac{1}{Z_k} \tag{5.134}$$

© 2007, first edition, EPFL Press

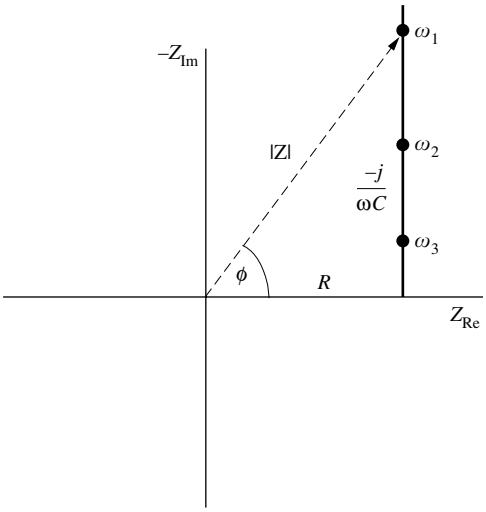


Figure 5.25 Impedance of electrical circuit consisting of a capacitance and a resistance in series.

The admittance of a circuit with elements connected in parallel is then:

$$A_\omega = \Sigma A_{\omega k} \tag{5.135}$$

If, for example, an electrical circuit is made of a resistance R and a capacitance C connected in series, its impedance is equal to the sum of the impedances of the two elements:

$$Z = Z_R + Z_C \tag{5.136}$$

With the values of Table 5.24, we obtain:

$$Z = R - \frac{j}{\omega C} \tag{5.137}$$

Figure 5.25 graphically represents this result in the complex plane. Equations (5.115) and (5.117) allow us to calculate the modulus and the phase shift.

$$|Z| = (R^2 + 1/\omega^2 C^2)^{1/2} \tag{5.138}$$

$$\phi = \arctan(-1/\omega CR) \tag{5.139}$$

Equivalent circuit of an electrochemical system

The theoretical interpretation of electrochemical impedance measurements must be built upon reaction models. With the equations of the model it is then possible to calculate the electrochemical impedance as a function of the frequency. A comparison

© 2007, first edition, EPFL Press

between theoretical and experimental impedances will then lead to the confirmation or rejection of the model. In many cases, it is useful to describe the impedance of an electrochemical system in terms of an electrical ***equivalent circuit*** made of passive elements. An equivalent circuit, of course, cannot replace a physical model, but it proves to be a practical tool that allows us to understand the electrical response of an electrode. The numerical fitting of impedance data with commercial software is generally based on the use of appropriate equivalent circuits.

In the simplest case, the equivalent circuit comprised of a capacitance C and a resistance R_t connected in parallel can describe the electrical behavior of the electrode-solution interface (Chapter 3). When a current flows, an ohmic resistance R_Ω must be added in series to take into account the ohmic drop in the electrolyte between the reference electrode and working electrode. Equations (5.141) and (5.142) express the impedance of the equivalent circuit presented in Figure 5.26. In these equations Z_C represents the impedance of the double layer,

$$Z_C = -\frac{j}{\omega C} \tag{5.140}$$

and R_t the charge-transfer resistance.

$$Z = R_\Omega + \frac{Z_C R_t}{Z_C + R_t} \tag{5.141}$$

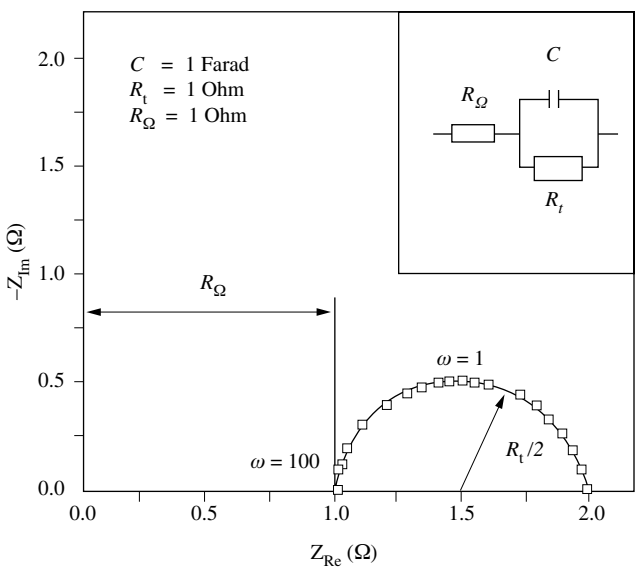


Figure 5.26 Nyquist diagram of impedance for a simple equivalent circuit of the electrode-electrolyte interface, calculated for the electrical circuit shown.

© 2007, first edition, EPFL Press

$$Z = R_\Omega - \frac{R_t\left(j/\omega C\right)}{R_t - j/\omega C} \tag{5.142}$$

In order to separate the different contributions in this result, the real and imaginary terms are rearranged by multiplying the second term by the complex conjugate of the denominator ($R_t + j/\omega C$):

$$Z = R_\Omega - j\left(R_t/\omega C\right)\frac{R_t + j/\omega C}{R_t^2 + 1/\omega^2 C^2}$$

$$= R_\Omega + \frac{R_t/\omega^2 C^2}{R_t^2 + 1/\omega^2 C^2} - j\frac{R_t^2/\omega C}{R_t^2 + 1/\omega^2 C^2} \tag{5.143}$$

The real part Z_{Re} and the imaginary part Z_{Im} of the impedance (5.143) are therefore equal to:

$$Z_{Re} = R_\Omega + \frac{R_t/\omega^2 C^2}{R_t^2 + 1/\omega^2 C^2} \tag{5.144}$$

$$Z_{Im} = -\frac{R_t^2/\omega C}{R_t^2 + 1/\omega^2 C^2} \tag{5.145}$$

The graphical representation of an impedance $Z = Z_{Re} + jZ_{Im}$ in the complex plane for different frequencies is called a ***Nyquist diagram***. Figure 5.26 shows an example of such a diagram, calculated using equations (5.144) and (5.145), by setting $R_\Omega = R_t = 1\ \Omega$ and $C = 1$ Farad. The impedance describes a semi-circle of radius $R_t/2$, whose center is located on the real axis at the distance $R_\Omega + 1/2\ R_t$ from the origin. In

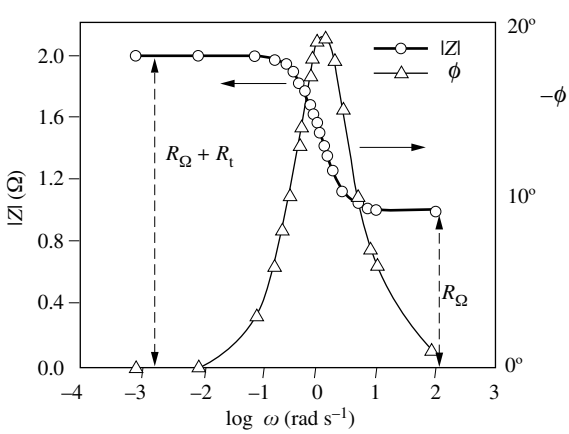


Figure 5.27 Bode diagram of impedance calculated for the electrical circuit of Fig. 5.26.

© 2007, first edition, EPFL Press

electrochemistry, it is customary to represent $-Z_{Im}$ on the ordinate (rather than Z_{Im}). The resulting semicircle is then inscribed in the first quadrant of the coordinate system.

By measuring the impedance of an electrode whose equivalent circuit corresponds to that shown in Figure 5.26, we can determine R_t and R_Ω, as well as the capacitance of the double layer C. The latter is obtained from the frequency at the maximum.

The impedance of an electrode can also be represented in a ***Bode diagram***, which displays the modulus $|Z|$ and the phase shift ϕ as a function of the logarithm of the radial frequency ω. Figure 5.27 shows the Bode diagram for the circuit drawn in Figure 5.26 using the same values of R_t, R_Ω and C. Equations (5.115) and (5.117) with (5.144) and (5.145) give the modulus and the phase shift.

Faradaic impedance

The reaction mechanism corresponding to the equivalent circuit shown in Figure 5.26 does not take mass transport or adsorption phenomena into account and therefore provides an incomplete picture of electrode reactions. Figure 5.28 shows a more realistic equivalent circuit. It still includes the ohmic resistance R_Ω and the double layer capacity C. However, the transfer resistance R_t is replaced by the ***faradaic impedance*** Z_F, which may contain one or more circuit elements, in series or in parallel. In order to evaluate the faradaic impedance, one needs a physical reaction model.

When transport phenomena play a role, the faradaic impedance can be separated into two terms, the kinetic impedance Z_t and the diffusion impedance Z_d.

$$Z_F = Z_t + Z_d \tag{5.146}$$

The ***kinetic impedance*** Z_t represents the faradaic impedance in the absence of a concentration overpotential. In the simplest case, the kinetic impedance corresponds to the transfer resistance R_t, but in more complicated situations it may include several circuit elements. The diffusion impedance Z_d describes the contribution of the concentration overpotential to the faradaic impedance and therefore depends on the transport phenomena in solution. In the absence of convection, it is referred to as the ***Warburg impedance*** Z_W and, in the opposite case, as the ***Nernst impedance*** Z_N.

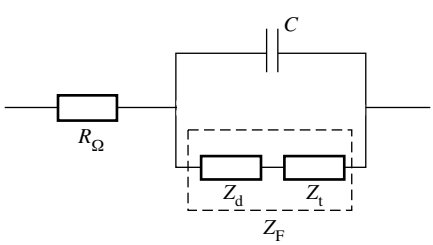


Figure 5.28 Randles equivalent circuit.

© 2007, first edition, EPFL Press

Neither of the two can be represented by a simple combination of passive elements of an electrical circuit.

Warburg Impedance

We derive an expression for the Warburg impedance for an electrode reaction of type

$$M^{n+} + n\,e = M \tag{5.147}$$

under the assumption that charge transfer is rapid. The electrode potential, at all times, is therefore given by the Nernst equation,

$$E = E^0 + \frac{RT}{nF} \ln c_s \tag{5.148}$$

where c_s represents the concentration of the M^{n+} ions at the electrode surface (Section 4.3).

For the following discussion we use the subscript *st* to express steady state conditions. Thus the steady state potential is E_{st}, the corresponding surface concentration is $c_{st,s}$ and the Nernst equation at steady state is written:

$$E_{st} = E^0 + \frac{RT}{nF} \ln c_{st,s} \tag{5.149}$$

We now superimpose a sinusoidal modulation of small amplitude, $\Delta E = |\Delta E| \exp(j\omega t)$, where $\Delta E = E - E_{st}$. Equation (5.150) expresses the concentration overpotential resulting from the modulation of the potential.

$$\Delta E = \frac{RT}{nF} \ln \frac{c_s}{c_{st,s}} = \frac{RT}{nF} \ln \left(\frac{\Delta c_s}{c_{st,s}} + 1 \right) \tag{5.150}$$

For small concentration variations this equation can be linearized:

$$\Delta E = \frac{RT}{nF} \frac{\Delta c_s}{c_{st,s}} \tag{5.151}$$

From equation (5.151) we deduce that the concentration at the electrode surface varies in the same fashion as the potential; a sinusoidal variation of potential therefore induces a sinusoidal variation of concentration near the electrode, given by

$$\Delta c = \frac{nF c_{st,s}}{RT} |\Delta E| \exp(j\omega t) \tag{5.152}$$

In order to calculate the Warburg impedance, we must first solve Fick's second law, which describes the concentration change $c - c_{st}$ induced by the potential modulation in the vicinity of the electrode.

© 2007, first edition, EPFL Press

$$\frac{\partial \Delta c}{\partial t} = D \frac{\partial^2 \Delta c}{\partial y^2} \tag{5.153}$$

The time derivative of equation (5.152) with (5.151) yields the following differential equation,

$$j\omega \Delta c = D \frac{\partial^2 \Delta c}{\partial y^2} \tag{5.154}$$

which has the general solution:

$$\Delta c = C_1 \exp\left(y (j\omega / D)^{1/2}\right) + C_2 \exp\left(-y (j\omega / D)^{1/2}\right) \tag{5.155}$$

In this equation, C_1 and C_2 are constants. To calculate the Warburg impedance, two boundary conditions are required:

$$\text{at } y = \infty, \quad \Delta c = 0 \tag{5.156}$$

$$\text{at } y = 0, \quad \Delta I = -n F A D \frac{\partial \Delta c}{\partial y} \tag{5.157}$$

The first condition stipulates that the modulation of the concentration only affects a limited zone in the vicinity of the electrode. As a result, $C_1 = 0$. Boundary condition (5.157) indicates that the current, at all times, is proportional to the concentration gradient at the electrode surface. With (5.155) we obtain:

$$\Delta I = -n F A D \left.\frac{\partial \Delta c}{\partial y}\right|_{y=0} = C_2 n F A D (j\omega / D)^{1/2} \tag{5.158}$$

$$C_2 = \frac{\Delta I}{n F A D (j\omega / D)^{1/2}} \tag{5.159}$$

The concentration variation at the electrode surface *(y = 0)* is therefore given by:

$$\Delta c_s = C_2 = \frac{\Delta I}{n F A D (j\omega / D)^{1/2}} \tag{5.160}$$

Equations (5.151) and (5.160) provide a relation between the variation of the potential and the current.

$$c_{st,s} \frac{nF}{RT} \Delta E = \frac{\Delta I}{n F A (j\omega D)^{1/2}} \tag{5.161}$$

© 2007, first edition, EPFL Press

The Warburg impedance is thus equal to:

$$Z_W = \frac{\Delta E}{\Delta I} = \frac{RT}{c_{\mathrm{st,s}}\, n^2 F^2 A\, (j\omega D)^{1/2}} \tag{5.162}$$

The mathematical identity

$$\frac{1}{\sqrt{j}} = \pm\left[\frac{1}{\sqrt{2}} - \frac{j}{\sqrt{2}}\right] \tag{5.163}$$

allows us to transform (5.162) by taking into account the fact that the real part must be positive:

$$Z_W = \frac{\sigma}{\omega^{1/2}} - j\,\frac{\sigma}{\omega^{1/2}} \tag{5.164}$$

$$\sigma = \frac{RT}{c_{\mathrm{st}}\, n^2 F^2 A\, (2D)^{1/2}}$$

With (5.116) we deduce from this relation:

$$\tan\phi = \frac{Z_{\mathrm{Im}}}{Z_{\mathrm{Re}}} = -1 \tag{5.165}$$

The phase shift is thus equal to $\phi = -45°$. In the Nyquist diagram, the Warburg impedance is therefore represented by a straight line at a 45° angle.

Nernst impedance

A derivation analogous to the previous yields the Nernst impedance, once the boundary conditions are modified to include the effects of convection: $y = \delta$, $\Delta c = 0$, where δ represents the thickness of the Nernst diffusion layer. The solution of (5.153) gives, in this case:

$$Z_N = \frac{\sigma}{\omega^{1/2}}\,(B+C) - j\,\frac{\sigma}{\omega^{1/2}}\,(B-C) \tag{5.166}$$

with

$$B = \sinh\lambda/(\cosh\lambda + \cos\lambda)$$

$$C = \sin\lambda/(\cosh\lambda + \cos\lambda)$$

$$\lambda = \delta\,(2\omega/D)^{1/2}$$

$$\sigma = RT/c_{st,s}\; n^2 F^2 A\; (2D)^{1/2}$$

© 2007, first edition, EPFL Press

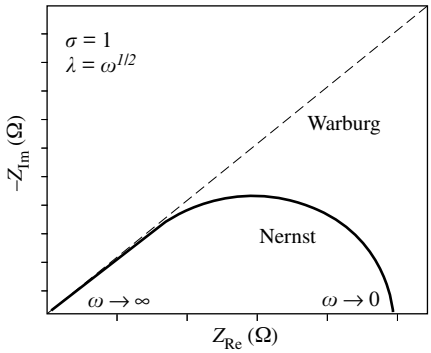


Figure 5.29 Nyquist diagram showing Warburg and Nernst impedances calculated for $\sigma = 1\ \Omega\ \mathrm{s}^{1/2}$ and $\lambda = \omega^{1/2}$.

At high frequency, the Nernst impedance becomes equal to the Warburg impedance since for $\lambda = \infty$: $B = 1$, $C = 0$. Figure 5.29 represents Z_W and Z_N in the complex plane. The Nernst impedance goes through a maximum then drops back down to the real axis at low frequencies.

Randles equivalent circuit.

The Warburg and Nernst impedances were derived under the assumption that the potential obeys the Nernst equation. The more realistic Randles model takes into account the kinetics of charge transfer as described by the Butler-Volmer equation. For the electrode reaction (5.147) this is written as

$$I = k'_{a(E)} - c_s\, k'_{c(E)} \tag{5.167}$$

with

$$k'_{a(E)} = n\,FA\, k_a \exp(E/\beta_a)$$

$$k'_{c(E)} = n\,FA\, k_c \exp(-E/\beta_c)$$

The small amplitude of the perturbation of the potential allows us to linearize relation (5.167), simplifying at the same time the expression describing the variation of the current.

$$\begin{aligned} \Delta I &= \left(\frac{\partial I}{\partial E}\right)_c \Delta E + \left(\frac{\partial I}{\partial c_s}\right)_E \Delta c_s \\ &= \left(\frac{k'_{a(E)}}{\beta_a} + k'_{c(E)}\frac{c_s}{\beta_c}\right)\Delta E - k'_{c(E)}\Delta c_s \end{aligned} \tag{5.168}$$

© 2007, first edition, EPFL Press

By setting

$$a' = \frac{k'_{\mathrm{a(E)}}}{\beta_{\mathrm{a}}} + k'_{\mathrm{c(E)}} \frac{c_{\mathrm{s}}}{\beta_{\mathrm{c}}} \tag{5.169}$$

we get:

$$\frac{\Delta I}{\Delta E} = a' - k'_{\mathrm{c(E)}} \frac{\Delta c}{\Delta E} = a' - k'_{\mathrm{c(E)}} \frac{\Delta c_{\mathrm{s}}}{\Delta I} \frac{\Delta I}{\Delta E} \tag{5.170}$$

With (5.160), this becomes:

$$Z = \frac{1}{a'} \left(1 + \frac{k'_{\mathrm{c(E)}}}{n F A (j \omega D)^{1/2}} \right) \tag{5.171}$$

The identity (5.163) allows us to separate the real and imaginary parts:

$$Z = \frac{1}{a'} + \frac{k'_{\mathrm{c(E)}}}{n F A a' (2 \omega D)^{1/2}} - j \frac{k'_{\mathrm{c(E)}}}{n F A a' (2 \omega D)^{1/2}} \tag{5.172}$$

The first term of the real part, which does not depend on the frequency, represents the transfer resistance: $R_{\mathrm{t}} = 1/a'$. By defining

$$\kappa' = \frac{k'_{\mathrm{c(E)}} R_{\mathrm{t}}}{n F A (2 D)^{1/2}} \tag{5.173}$$

we get the faradaic impedance of an electrode, composed of a transfer resistance R_{t} and a Warburg resistance.

$$Z_{\mathrm{F}} = R_{\mathrm{t}} + Z_{\mathrm{W}} = R_{\mathrm{t}} + \kappa' \omega^{-1/2} - j \kappa' \omega^{-1/2} \tag{5.174}$$

The total impedance of the electrode corresponds to that of the ***Randles equivalent circuit***, shown in Figure 5.30. If R_{Ω} represents the ohmic resistance between the working and reference electrodes, and if C designates the double layer capacity, the impedance of this circuit is expressed by the function:

$$Z = R_{\Omega} + Z_{\mathrm{F}} \frac{-j / \omega C}{Z_{\mathrm{F}} - j / \omega C} \tag{5.175}$$

By replacing Z_{F} by (5.174) we obtain:

$$Z = R_{\Omega} + \frac{R_{\mathrm{t}} + \kappa' \omega^{-1/2}}{M} - j \frac{\kappa' \omega^{-1/2} + \omega C \left(\kappa' \omega^{-1/2} \right)^2 + \omega C \left[R_{\mathrm{t}} + \left(\kappa' \omega^{-1/2} \right) \right]^2}{M} \tag{5.176}$$

© 2007, first edition, EPFL Press

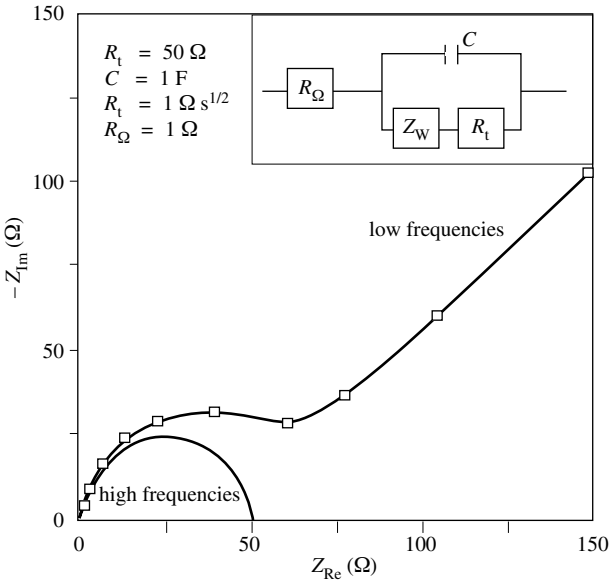


Figure 5.30 Nyquist diagram of impedance calculated for Randles circuit as shown.

with

$$M = \omega^2 C^2 \left[R_t + \left(\kappa' \omega^{-1/2} \right) \right]^2 + \omega^2 C^2 \left[\kappa' \omega^{-1/2} + \left(1 / \omega C \right) \right]^2 \tag{5.177}$$

In Figure 5.30, the Nyquist diagram consists of two distinct domains: At high frequency, a semicircle shows that the transfer resistance determines the impedance, while at low frequencies, diffusion related phenomena dominate, and a Warburg impedance is observed.

Application of impedance measurements in corrosion

In general, the impedance of solid electrodes exhibits a more complicated behavior than predicted by the Randles model. Several factors are responsible for this. Firstly, the simple Randles model does not take into account the time constants of adsorption phenomena and the individual reaction steps of the overall charge transfer reaction (Section 5.1). In fact the kinetic impedance may include several time constants, and sometimes one even observes inductive behavior. Secondly, surface roughness or non-uniformly distributed reaction sites lead to a dispersion of the capacitive time constants. As a consequence, in a Nyquist plot the semicircle corresponding to a charge-transfer resistance in parallel to the double-layer capacitance becomes flattened. To account for this effect it has become current practice in corrosion science and engineering to replace the double layer capacitance in the equivalent circuit by a

© 2007, first edition, EPFL Press

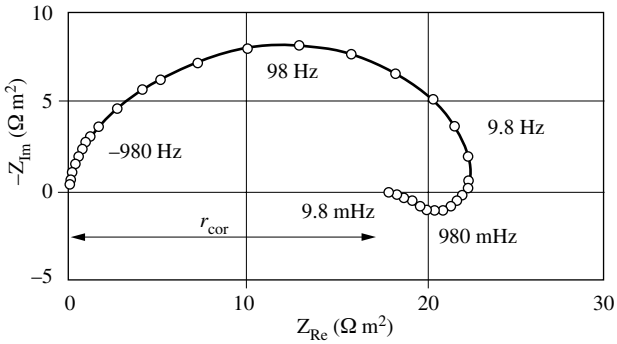


Figure 5.31 Nyquist diagram of the impedance of an iron electrode in daerated 0.5 M H_2SO_4 [13].

constant phase element or CPE [12]. The admittance of a constant phase element is given by

$$A_{CPE} = A_0 (j\omega)^{\alpha} \tag{5.177}$$

where the admittance $A_{CPE} = 1/Z_{CPE}$ corresponds to the reciprocal of the impedance, A_0 is the admittance of the CPE at $\omega = 1$ rad/s and α is a coefficient. For $\alpha = 1$ the frequency dependence of the CPE is that of a capacitance. To fit a flattened hemicircle the value of α is usually taken as an adjustable parameter and its value is typically slightly lower than one. While the constant phase element approach permits to fit an

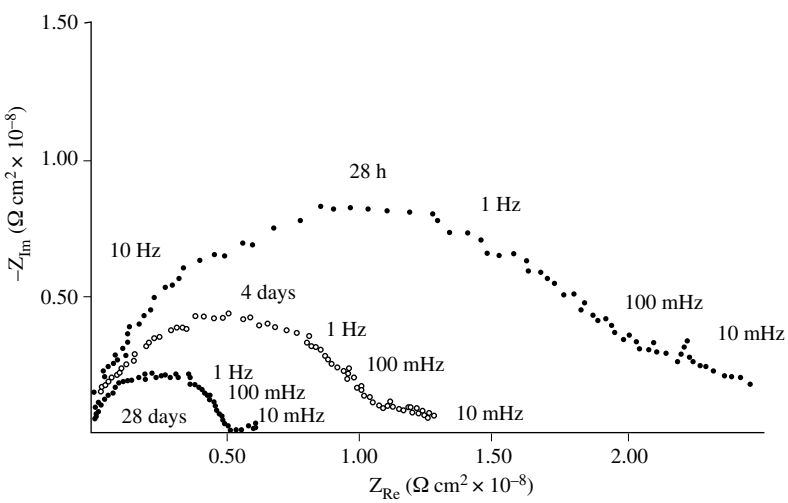


Figure 5.32 Nyquist diagram showing the impedance of paint-coated steel in NaCl at different exposure times [14].

© 2007, first edition, EPFL Press

impedance diagrams showing a depressed hemicircle, we must keep in mind that in most cases it lacks a clear physical basis.

The Nyquist diagram of Figure 5.31, measured on steel in sulfuric acid, exhibits a flattened hemicircle and an inductive loop at low frequency. It illustrates that even apparently simple corrosion systems often exhibit complicated impedance behavior. The Figure indicates how to determine the rate of uniform corrosion from impedance measurements. For this one determines the polarization resistance at the corrosion potential, r_{cor}, from the Nyquist diagram as shown in the figure. Equation (4.73) then permits to calculate the corrosion rate, provided the Tafel coefficients are known. In contrast to steady state polarization measurements, impedance measurements allow us to easily separate the ohmic resistance R_Ω from the polarization resistance (Figure 5.26). The impedance technique is therefore well suited for measuring corrosion rates in poorly conducting media, or of metals covered by a high-resistance surface layer, such as an organic coating.

Figure 5.32 illustrates the application of impedance spectroscopy to the investigation of the rate of deterioration of an organic coating. It shows the Nyquist diagram of an iron sample, coated with an acrylic paint, and then immersed into an NaCl solution for 28 hours, 4 days, and 28 days, respectively. The diameter of the flattened semicircle decreases with exposure time. The observed decrease in the impedance is a measure of the progressive degradation of the coating.

While it is often difficult to deduce a unique reaction mechanism from impedance measurements, a detailed knowledge of mechanism is not always required. Therefore, the impedance method finds many practical applications in corrosion science and engineering; it is used, for example, in the measurement of corrosion rates, the development of organic coatings, and the characterization of surface films.

Bibliography

K. J. Vetter, *Electrochemical Kinetics*, Academic Press, New York (1967) 788 pp.

J. O'M Bockris, A. K. N. Reddy, *Modern Electrochemistry*, vol. 2, Plenum Press, New York (1970) pp. 623-1432.

A. J. Bard, L. R. Faulkner, *Electrochemical Methods*, 2nd Edition, John Wiley, New York (2001) 836 pp.

References

[1] E. Mattson, J. O'M. Bockris, *Trans. Farad. Soc.* 55, 1586 (1959).

[2] K. E. Heusler, *Encyclopedia of Electrochemistry of the Elements*, vol. 9, part A, A. J. Bard, ed., Marcel Dekker Inc., New York (1982) pp. 230-360.

[3] K. J. Vetter, *Electrochemical Kinetics,* Academic Press, New York (1967) p. 539.

[4] R. Greef, R. Peat, L. M. Peter, D. Pletcher, J. Robinson, *Instrumental Methods in Electrochemistry*, Ellis Horwood Ltd., Chichester (1985) p. 236.

[5] Ref. 3, p. 657.

[6] H. C. Kuo, D. Landolt, *Corr. Sci.* 16, 915 (1976).

© 2007, first edition, EPFL Press

[7] J. R. Selman, C. W. Tobias, *J. Electroanal. Chem.*, 65, 67 (1975).
[8] R. S. Nicholson, I. Shain, *Anal. Chem.* 36, 706 (1964).
[9] M. R. Tarasevich, A. Sadowski, E. Yeager, *Treatise of Electrochemistry,* vol. 7, Plenum Press, New York (1983) p. 314.
[10] C. Gabrielli, *Identification of Electrochemical Processes by Frequency Response Analysis*, Technical Report 004/83, Solartron Analytical, Paris (1983) pp. 1-128.
[11] M. Sluyters-Rehbach, J. H. Sluyters, *Electroanalytical Chemistry*, vol. 4, A. J. Bard, ed., Marcel Dekker Inc., New York (1970) p. 3.
[12] F. Mansfeld, "Electrochemical Impedance Spectroscopy", in *Analytical Methods in Corrosion Science and Engineering*, P. Marcus and F. Mansfeld, eds., CRC Press, Boca Raton 2005, 463.
[13] W. J. Lorenz, F. Mansfeld, *Electrochemical Corrosion Testing*, E. Heitz , J. C. Rowland, F. Mansfeld, eds., Verlag Chemie, Weinheim (1986) p. 185.
[14] J. Hubrecht, J. Vereecken, M. Piens, *J. Electrochem. Soc.* 131, 2010 (1984).

© 2007, first edition, EPFL Press

CHAPTER 6

PASSIVITY OF METALS

6.1 TRANSITION FROM ACTIVE TO PASSIVE STATE

6.1.1 Active and passive metals

Metals in the ***passive state*** (passive metals) have a thin oxide layer on their surface, the ***passive film,*** which separates the metal from its environment. Metals in the ***active state*** (active metals) are film free. Most metals and alloys that resist well against corrosion are in the passive state: stainless steel, nickel-chromium based superalloys, titanium, tantalum, aluminum, etc. Typically, the thickness of passive films formed on these metals is about 1–3 nm.

Dissolution of an active metal (***active dissolution***), involves a charge transfer at the metal-electrolyte interface. Soluble ions, either hydrated or complexed, are formed and dissolve into the electrolyte, while the liberated electrons either flow to the cathode or are taken up by an oxidizing agent.

$$\mathrm{M} \rightarrow \mathrm{M}^{n+}_{(aq)} + n\,e \tag{6.1}$$

The reaction kinetics of active metal dissolution obeys the Butler-Volmer equation and therefore, the anodic current density increases monotonically with the potential.

When a passive metal dissolves (***passive dissolution***), cations are formed by a charge transfer reaction at the metal-film interface. They migrate across the passive film to the film-electrolyte interface, where they dissolve into solution as hydrated or complexed ions.

$$\mathrm{M} + \frac{n}{2}\mathrm{H_2O} \rightarrow \mathrm{MO}_{n/2} + n\,\mathrm{H}^+ + n\,e \tag{6.2}$$

$$\underline{\mathrm{MO}_{n/2} + n\,\mathrm{H}^+ \rightarrow \mathrm{M}^{n+}_{(aq)} + \frac{n}{2}\mathrm{H_2O}} \tag{6.3}$$

$$\mathrm{M} \rightarrow \mathrm{M}^{n+}_{(aq)} + n\,e \tag{6.1}$$

© 2007, first edition, EPFL Press

The overall stoichiometric equation (6.1) is the same for active and for passive dissolution, but depending on the metal the charge number, *n*, may differ between the two. Examples are presented in Table 6.1.

Table 6.1 Ionic species formed by active and passive dissolution of iron, nickel, and chromium in acid media.

Metal	Active dissolution	Passive dissolution
iron	Fe^{2+}	Fe^{3+}
nickel	Ni^{2+}	Ni^{2+}
chromium	Cr^{2+}	Cr^{3+}

Because of the presence of an oxide film, the dissolution rate of a passive metal at a given potential is much lower than that of an active metal. It depends mostly on the properties of the passive film and its solubility in the electrolyte. During ***passivation***, which is a term used to describe the transition from the active to the passive state, the rate of dissolution therefore decreases abruptly. The polarization curve of a stainless steel in sulfuric acid, given in Figure 6.2, illustrates this phenomenon. In this electrolyte, the corrosion potential of the alloy is close to –0.3 V. Anodic polarization leads to active dissolution up to about –0.15 V, where the current density reaches a maximum. Beyond this point, the current density, and hence the dissolution rate, drops sharply. It then shows little further variation with potential up to about 1.1 V. Above that value the current density increases again because transpassive dissolution and oxidation of water to oxygen becomes possible.

$$2\,H_2O \rightarrow O_2 + 4\,H^+ + 4\,e \tag{6.4}$$

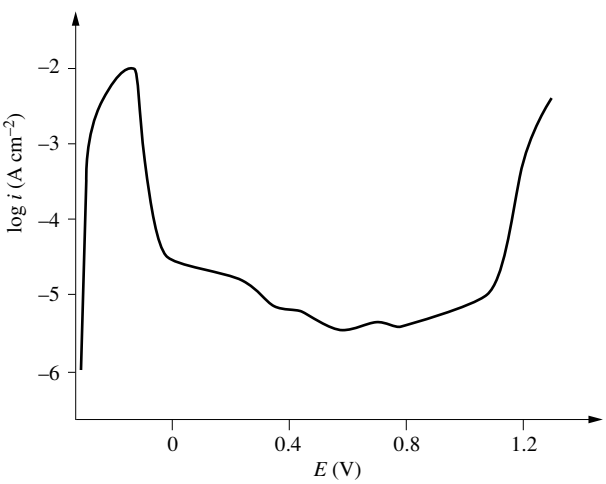


Figure 6.2 Polarization curve of Fe-17Cr stainless steel in 0.5 M H_2SO_4. Sweep rate is 0.02 V/min. [1].

© 2007, first edition, EPFL Press

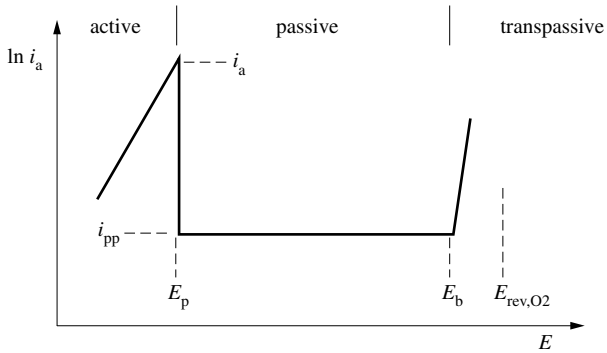


Figure 6.3 Variation of partial anodic current density with potential for a passivating metal.

Generally speaking, we can distinguish three potential regions in the polarization curve of a passivating metal (Figure 6.3):

- the active region;
- the passive region;
- the transpassive region.

In contrast to the active and passive regions, the surface state of the metal in the transpassive region is not well defined and an oxide may or may not cover the surface.

Definitions

The current density measured during a polarization experiment is the sum of all anodic and cathodic partial current densities (Chapter 4). Figure 6.3 schematically shows the variation of the anodic partial current density of a passivating metal as a function of the potential. It allows us to define a number of quantities that describe the polarization behavior of passivating metals.

The ***passivation potential*** E_p separates the active from the passive potential region. The corresponding current density at the maximum is the ***passivation current density***, i_p. The ***passive current density*** i_{pp} characterizes the dissolution behavior of the metal in the passive potential region. The ***transpassivation potential*** E_b marks the end of the passive potential region and the transition from passive to transpassive behavior. Beyond this point the anodic partial current density increases markedly with increasing potential due one of the following processes: uniform transpassive dissolution resulting from oxidation of the passive film, dissolution by pitting resulting from local film breakdown, oxygen evolution due to water oxidation. When dissolution by pitting is the dominating reaction, the potential E_b is called ***pitting potential*** or *critical pitting potential*. Often, E_b is also referred to as ***film breakdown potential***, indicating that pitting is initiated by passive film breakdown (see Section 6.3). Depending on conditions, the value of E_b can be either above or below the reversible potential of the oxygen electrode, $E_{rev,O2}$. For sufficiently stable passive films with

© 2007, first edition, EPFL Press

good electronic conductivity, oxygen evolution rather than transpassive dissolution may therefore account for the observed current at high anodic potentials.

The pitting potential, plays an important role for the corrosion resistance of passive metals and alloys as we shall see in more detail in Chapter 7. Generally speaking, to have a good corrosion resistance an alloy should exhibit a low value of E_p and a high value of E_b.

6.1.2 Thermodynamic and kinetic aspects of passivation

Anodic passivation and spontaneous passivation

For a metal to become passive its potential must exceed the passivation potential, a condition that can be reached in two ways:

- by anodic polarization (anodic passivation);
- by reaction with an oxidizing agent (spontaneous passivation).

The theory of mixed electrodes allows us to see that these two modes of passivation are not fundamentally different. Figure 6.4 presents the Evans diagram of a passivating metal, exposed to (a) a weakly oxidizing environment; and (b) a strongly oxidizing environment.

In the weakly oxidizing solution (a), the corrosion potential is located in the active potential region ($E_{cor} < E_p$). In order to passivate the metal, an anodic potential of $E > E_p$ has to be applied by connecting the metal as the anode in an electrochemical cell. We then speak of ***anodic passivation***. An anodically passivated metal generally loses its passivity as soon as the electrical circuit is turned off. In the strongly oxidizing environment (b) of Figure 6.4, the oxidizing agent causes a ***spontaneous passivation***, also called *chemical passivation*, of the metal. The corrosion potential in this case is located in the passive region: $E_p < E_{cor} < E_b$.

Figure 6.5 illustrates how the presence of an oxidizing agent affects the observed polarization behavior of a metal. It shows the polarization curves of a nickel-chromium-molybdenum alloy in sulfuric acid, with and without ferric sulfate added

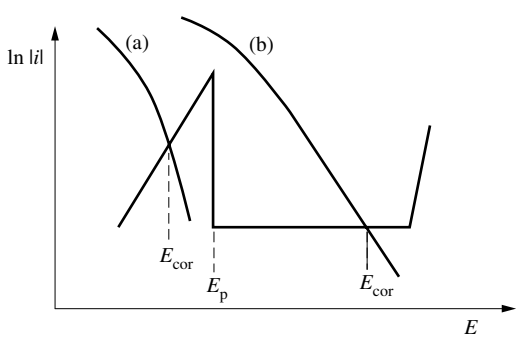


Figure 6.4 Effect of cathodic partial reaction on corrosion potential: (a) slightly oxidizing environment, $E_{cor} < E_p$; (b) oxidizing environment, $E_{cor} > E_p$.

© 2007, first edition, EPFL Press

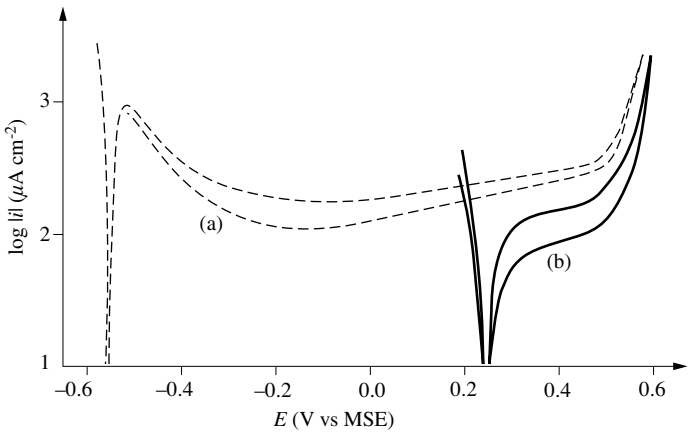


Figure 6.5 Anodic polarization curve of Fe-66Ni-16Mo in 50% H_2SO_4 in absence (a) and in presence (b) of 42g $Fe_2(SO_4)_3$; Sweep rate is 0.33 mV/s, $T = 125$ °C [2].

as oxidizing agent. In the absence of the ferric sulfate, the corrosion potential is located in the active region, whereas in presence of ferric sulfate it is located in the passive region. We may note that the active-passive transition in the polarization curve of the spontaneously passive metal is hidden by the cathodic partial current for the reduction of the oxidizing agent.

Under certain conditions, the anodic and cathodic partial current densities may intersect at two points, as shown in Figure 6.6. In this case, the corrosion potential can be located in either the active or the passive region, depending on the preceding surface treatment. A surface that has been activated beforehand (for example, by pickling in hydrochloric acid), will remain active, while a surface that has been previously passivated will maintain its passive state, but only as long as the oxide film is not damaged mechanically, for example by abrasion or scratching. The described situation is referred to as ***unstable passivity***.

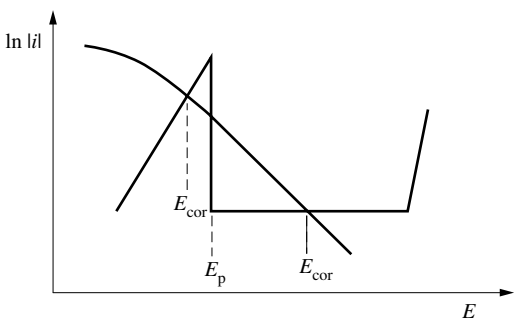


Figure 6.6 Evans diagram showing a situation of unstable passivity. The corrosion potential can be located in either the active or passive region.

© 2007, first edition, EPFL Press

The behavior of iron in nitric acid illustrates this phenomenon. In concentrated nitric acid the iron is passive, while in dilute nitric acid it remains active. The behavior in a nitric acid solution of intermediate concentration depends on the pretreatment of the sample: first dipped into dilute nitric acid, the iron will retain its active state in the solution of intermediate strength, while first dipped in concentrated acid it will be passive in the same solution. The passivity will be lost, however, as soon as a scratch damages the passive film.

Conditions for spontaneous passivation

Generally, a metal that passivates spontaneously in a given environment, will resist corrosion better than a metal exhibiting a corrosion potential in the active region. It is therefore important to study the kinetic and thermodynamic parameters that control spontaneous passivation of a metal. For passivation to be spontaneous and stable, the two following conditions must be fulfilled:

$$E_p < E_{rev,ox} \tag{6.5}$$

$$i_p < |\, i_{c(Ep)} \,| \tag{6.6}$$

Spontaneous passivation therefore depends on the following four parameters:

- the passivation potential, E_p;
- the reversible potential of the oxidizing agent, $E_{rev,ox}$;
- the passivation current density, i_p;
- the cathodic partial current density of the oxidizing agent at the passivation potential, $i_{c(Ep)}$.

The electrochemical properties of the metal (E_p, i_p) as well as those of the oxidizing agent ($E_{rev,ox}$, $i_{c(Ep)}$) are therefore important for the spontaneous passivation of metals.

Passivation potential

Figure 6.7 illustrates why a less noble passivation potential facilitates spontaneous passivation. It shows anodic partial current densities for two metals, whose only difference is the value of their passivation potential. The cathodic partial current densities are identical for the two metals. We observe that the corrosion potential of metal *a* having a lower passivation potential is located in the passive region, while that for metal *b* is in the active region. Metal *a*, although less noble, is therefore expected to better resist corrosion.

The passivation potential of a metal is equal or higher than the reversible potential of the electrode reaction (6.2), which describes oxide formation involving water.

$$E_p \geq E_{rev,oxide} \tag{6.5}$$

© 2007, first edition, EPFL Press

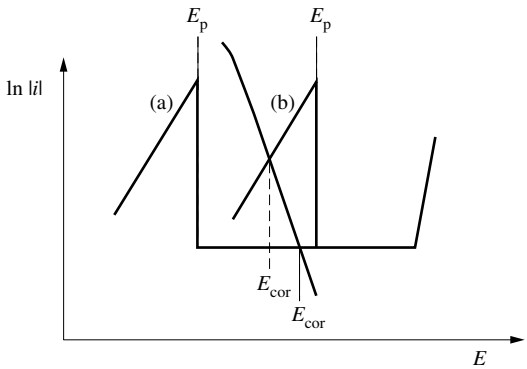


Figure 6.7 Evans diagram showing that, depending on the value of the passivation potential, the corrosion potential can be located in the passive (a) or active (b) region.

The reversible potential of reaction (6.2), $E_{rev,oxide}$, is equal to:

$$E_{rev,\,oxide} = E^\circ + \frac{RT}{nF} \ln (a_{H^+})^n = E^\circ + \frac{RT}{F} \ln a_{H^+} \tag{6.7}$$

By introducing the pH, we find for a temperature of 25 °C :

$$E_{rev,\,oxide} = E^\circ - 0.059\ pH \tag{6.8}$$

Table 6.8 Standard free energies and standard potentials of oxide formation (adapted from [3]).

Oxide	ΔG°/kJ mol^{-1}	E° / V
Ti_2O_3 (hydrated)	–1388	–1.169
Al_2O_3 (Gibbsite)	–1154.9	–0.766
Ta_2O_5	–1912.1	–0.753
Cr_2O_3	–1058	–0.599
ZnO	–318.3	–0.420
MoO_2	–533	–0.152
Fe_3O_4 (magnetite)	–1016.5	–0.087
Fe_2O_3 (hematite)	–742.2	–0.053
FeO (wustite)	–246.14	–0.041
NiO (unhydrated)	–216	0.110
MoO_3	–668	0.075
PbO (hydrated, hexagonal)	–183.7	0.277
CrO_3	–510	0.348
Pb_3O_4	–616.2	0.431
Cu_2O	–148.1	0.462
CuO	–134	0.535

© 2007, first edition, EPFL Press

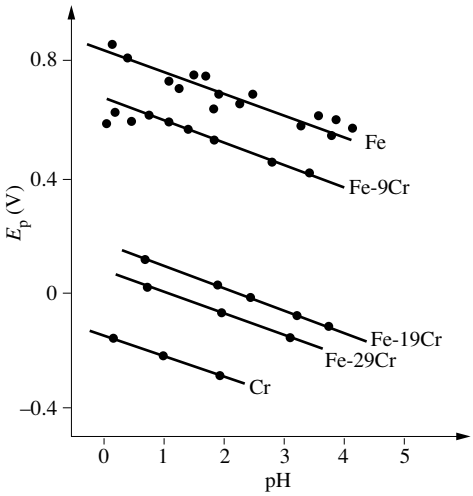


Figure 6.9 Variation of passivation potential for Fe-Cr alloys with pH in sulfate solutions [4].

Table 6.8 provides the Gibbs free energy of formation of different oxides and the corresponding standard potentials for aqueous solutions. Generally, the oxides of less noble metals exhibit a lower standard potential of formation. These metals passivate spontaneously in the presence of protons. Furthermore, many oxides exhibit good chemical stability in acidic environments. This explains the higher corrosion resistance of metals such as titanium, tantalum and chromium.

According to the relation (6.8), the value of $E_{rev,oxide}$ decreases by 59 mV per pH unit, regardless of the stoichiometry of the oxide formed; this is because the number of charges does not appear in the equation. The condition $E_p \geq E_{rev,oxide}$, suggests that the passivation potential also decreases with increasing pH. Figure 6.9 confirms this behavior: When the passivation potentials of iron, chromium and their alloys in sulfuric acid are plotted as a function of the pH, a straight line is obtained with a slope of –59 mV/pH [4]. Similar results have been found for other metals [5].

In this figure, for zero pH (pH = 0; $E_{rev,oxide} = E°$), the passivation potentials of chromium and iron do not match the standard potentials of the oxides Cr_2O_3 and Fe_2O_3 listed in Table 6.8. One explanation is that kinetic limitations lead to a higher passivation potential than predicted by thermodynamics. Another reason could be that the listed standard potentials were measured on bulk samples. The extreme thinness of passive films could influence their thermodynamic properties. In addition, their composition does not always correspond to a simple stoichiometry. For example, chromium-iron alloys form passive films containing both iron and chromium cations and their passivation (Figure 6.9) lie between those of iron and chromium. Then again, they exhibit the same pH dependence as the pure metals.

Passivation current density

Figure 6.10 illustrates, by means of an Evans diagram, the corrosion behaviour of two metals that differ only in terms of their passivation current densities. For metal *a*

© 2007, first edition, EPFL Press

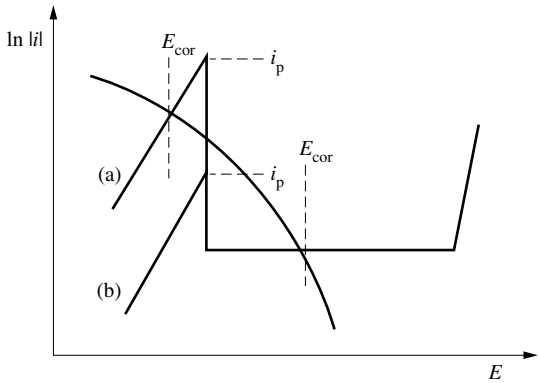


Figure 6.10 Evans diagram showing the effect of passivation current density on corrosion behavior: (a) corrosion potential in the active region; (b) corrosion potential in the passive region.

the passivation current density exceeds (in absolute value) the current density of the cathodic partial reaction at the passivation potential: $i_p > |i_{c(Ep)}|$. The corrosion potential in this case may be located in either the passive or the active region, and the passivity is unstable. Metal *b*, for which $i_p < |i_{c(Ep)}|$, exhibits a state of stable passivity and therefore a low rate of corrosion.

In acidic media, the metals iron, nickel and chromium have passivation current densities that increase in the order Cr < Ni < Fe. In Figure 6.11, the anodic polarization curves for the three metals in 0.5 M sulfuric acid (25 °C) are compared. Chromium has lower values of both i_p and E_p than the other two metals. By alloying increasing amounts of chromium to steel one therefore improves the corrosion resistance. Experience shows that above a chromium concentration of 12 to 13%, a steel passivates spontaneously in contact with aerated water. It becomes "stainless", meaning it does not rust easily. Figure 6.12 gives the corrosion potential of different

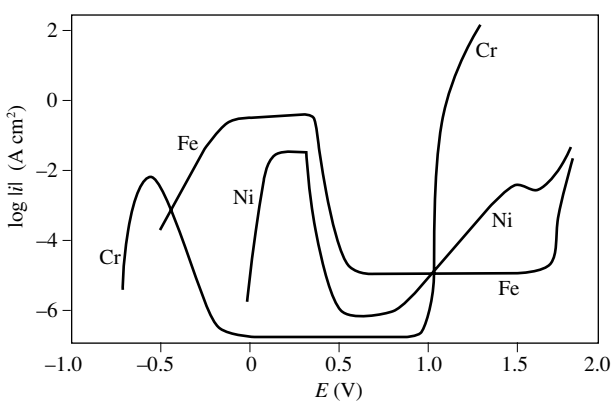


Figure 6.11 Anodic polarization curves of Fe, Ni and Cr in 0.5 M H_2SO_4 [6].

© 2007, first edition, EPFL Press

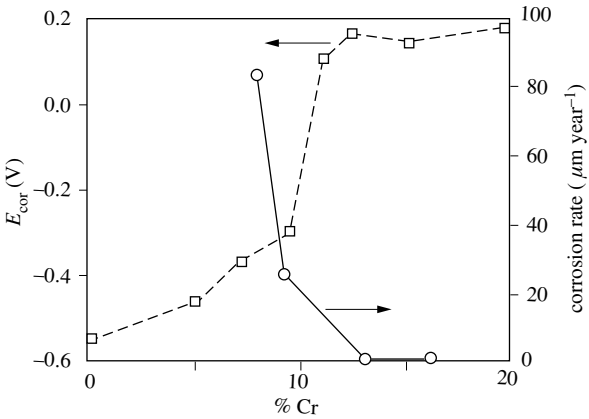


Figure 6.12 Corrosion potentials of Fe-Cr alloys in aerated 4% NaCl as a function of their chromium content. The measured average rate of corrosion in salt spray tests is also shown [7].

iron-chromium alloys in an aerated solution of 4% NaCl together with the corrosion rate measured in a salt spray test. Spontaneous passivation above a chromium content of 8 to 12%, leads to a rise in corrosion potential and to a drop of corrosion rate.

The magnitude of the passivation current density depends on different factors:

- the kinetics of active dissolution;
- the mass transport of the dissolution products;
- the pH of the electrolyte;
- the water content of the electrolyte.

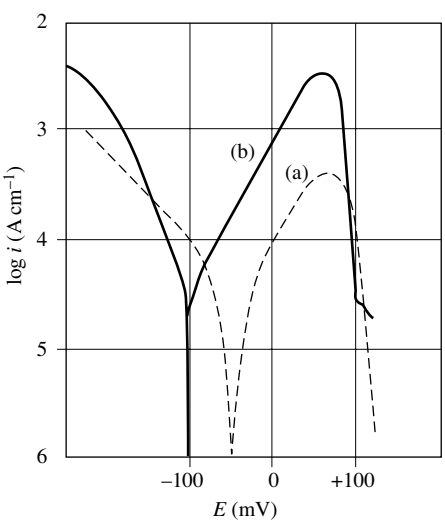


Figure 6.13 Polarization curves of Inconel 690 in 0.5 M H_2SO_4+1000 ppm Cl^-, in the absence (a) and in the presence (b) of 1 ppm dissolved H_2S. Sweep rate is 10 mV/min. [8].

© 2007, first edition, EPFL Press

The *exchange current density* influences the value of the passivation current density more than any other kinetic parameter. Indeed, if the anodic curves (a) and (b) of Figure 6.10 were to belong to the same metal, the difference in i_p could be explained by a lower exchange current density for case (b). Certain adsorbed species inhibit or catalyze the dissolution of metals and thus affect the effective value of the exchange current density. For example, sulfide ions catalyze the dissolution of nickel and its alloys by increasing the exchange current density. Figure 6.13 shows polarization curves of a nickel alloy in 0.5 M H_2SO_4, and in the same solution with 1ppm of H_2S added. The presence of H_2S shifts the anodic Tafel line to higher current densities, while it does not affect the passivation potential. As a result, the passivation current density in presence of H_2S is higher. Some alloying elements, such as the molybdenum added to certain stainless steels, inhibit anodic dissolution. The polarization curves plotted in Figure 6.14 serve as an illustration. They show that the passivation current density decreases as the proportion of molybdenum in a Fe-13Cr alloy increases. In addition, the presence of molybdenum lowers slightly the passivation potential.

Sometimes, the *mass transport* limits the passivation current density. An example is iron in sulfuric acid, where the passivation current density is equal to the anodic limiting current density for iron dissolution. The rate limiting step is the transport of ferrous ion from the electrode surface to the bulk solution. The passivation current density therefore is determined by the solubility of ferrous sulfate and by the prevailing

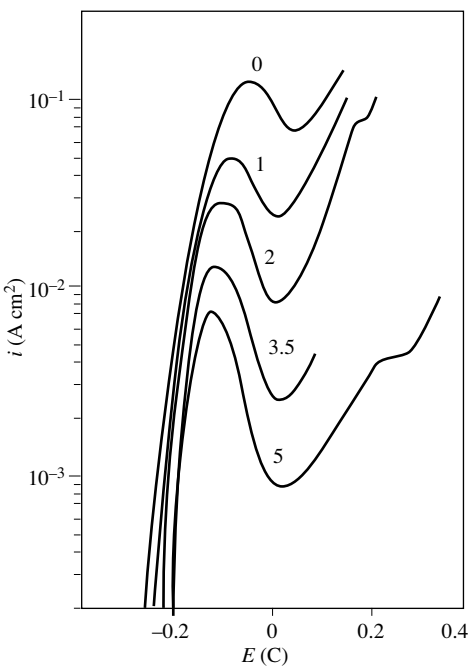


Figure 6.14 Polarization curves of Fe-13Cr-x Mo stainless steels in 1M HCl, showing the effect of alloyed molybdenum on the passivation current density. T = 30 °C, x = 0, 1, 2, 3.5, 5% [9].

© 2007, first edition, EPFL Press

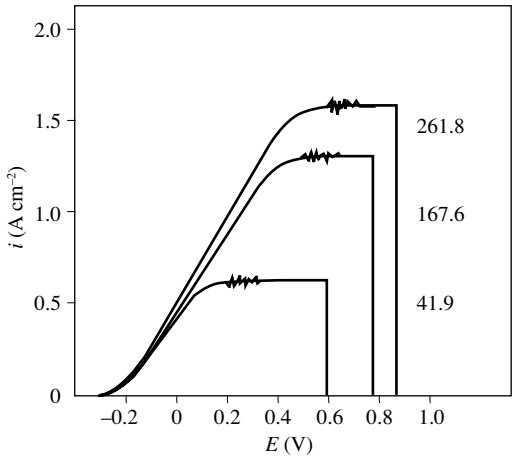


Figure 6.15 Anodic polarization curves of an Fe rotating disk electrode in 1 M H_2SO_4 at different rotation rate (rad s^{-1}) [10].

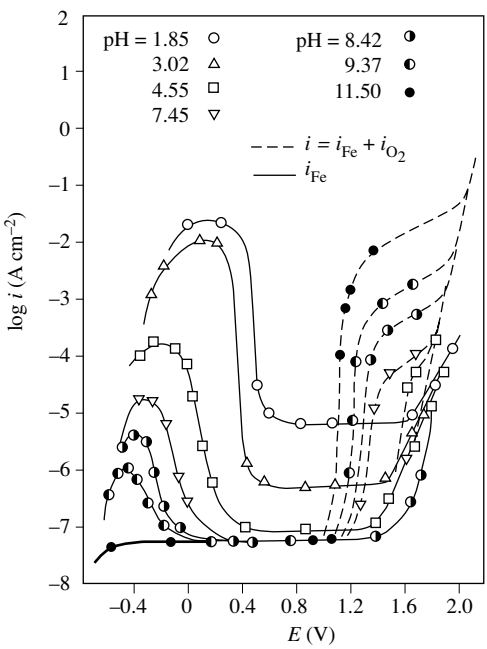


Figure 6.16 Anodic partial current density for the dissolution of iron measured in 0.15 M phosphate solutions of different pH. Dashed lines show the sum of the dissolution current densities and the formation of oxygen by the oxidation of OH^- ions and H_2O [11].

© 2007, first edition, EPFL Press

hydrodynamic conditions. The results shown in Figure 6.15, obtained with a rotating disk electrode, confirm the influence of hydrodynamics: with increasing rotation rate the passivation current density increases.

The *pH of the electrolyte* does not only have an effect on the passivation potential, but also on the passivation current density, because both the metal dissolution kinetics and the solubility of hydroxides depend on pH. Figure 6.16 shows that the passivation current density of iron becomes smaller at higher pH. This has been explained by a lowering of the solubility of ferrous hydroxide, which precipitates at the surface. Since both the passivation potential and the passivation current density decrease with increasing pH, spontaneous passivation of iron becomes possible in basic, aerated media. This explains why steel reinforcements in concrete (pH ≥ 13) resist corrosion well as long as chemical reactions with carbon dioxide from air (carbonation of concrete) do not modify the alkalinity.

According to equation (6.2), the anodic formation of oxides is the result of a reaction between the metal and water. Thus, the smaller the *activity of water*, the more difficult it becomes to passivate a metal. Figure 6.17 highlights this phenomenon: in an acetic acid/acetate/water mixture, the passivation current density decreases sharply when the proportion of water increases from 10 to 50%. If there is no water, no passive film is formed. Non-aqueous corrosive environments are encountered mostly in the chemical industry.

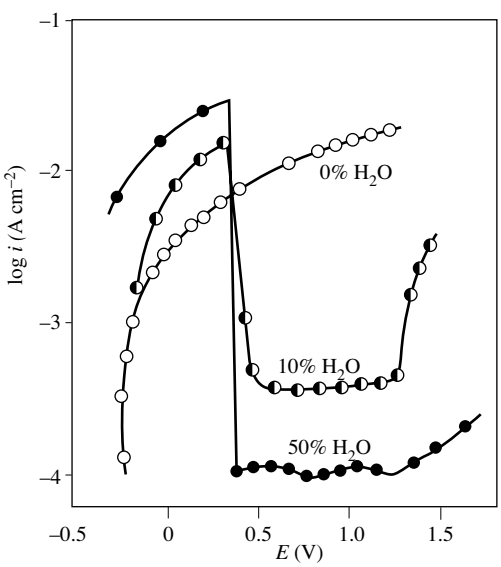


Figure 6.17 Influence of water content in the electrolyte on iron passivation in a mixture of acetic acid and sodium acetate [11].

© 2007, first edition, EPFL Press

6.2 GROWTH AND PROPERTIES OF PASSIVE OXIDE FILMS

6.2.1 Structure and composition of passive films

Experimental study of passive films

Because passive films are extremely thin (1–3 nm), the usual techniques of chemical analysis and X-ray diffraction are not suited for determining their composition and structure. Furthermore, exposed to an electrolytic environment, the passive film is continually regenerating itself, with the dissolution rate of the film equal to the corrosion rate of the metal. Indeed, this capacity of self-regeneration is an important property of passive films. Despite their immense technological importance, our knowledge of the structure, composition and properties of passive films is therefore still limited.

For the study of passive films, *in-situ* methods (study of the film in solution) and *ex-situ* techniques (study of the film after removal from the electrolyte) are available. In principle, in-situ techniques are preferable because they allow us to avoid any chemical modification that might emerge during the transfer of the sample from the electrolyte to the instrument used for analysis. However, because ex-situ techniques, in particular the surface analysis methods discussed in Chapter 3, give more complete chemical information, they have been most widely used.

Among the *in-situ* techniques, ***coulometry*** is a classical method for the determination of the amount of oxide present on a surface, and in certain cases, of its oxidation state. The technique consists of measuring the charge required to reduce the passive film when a weak cathodic current density (on the order of μA cm^{-2}) is applied to the electrode. For example, the cathodic reduction of Fe_2O_3 produces dissolved ferrous ions:

$$Fe_2O_3 + 6\,H^+ \rightarrow 2\,Fe^{2+} + 3H_2O + 2\,e \tag{6.9}$$

The potential recorded as a function of time yields a transient similar to the one shown in Figure 6.18. A potential plateau is observed, which corresponds to the reduction of the oxide. In the absence of important kinetic effects, the plateau potential is determined by the reversible potential of oxide formation and therefore, in principle, permits to identify the oxide stoichiometry. The charge for oxide reduction is proportional to the length of the plateau. This permits to estimate the film thickness provided one knows the reaction stoichiometry and the density of the film.

The described technique is applicable to chemically stable passive films that can be cathodically reduced. Some passive films dissolve spontaneously when the anodic polarization is interrupted. In that case, if one follows the evolution of the corrosion potential as a function of time one observes a transient that resembles that shown in Figure 6.18. The potential of the plateau, called the ***Flade potential***, is indicative of the passivation potential. However, the length of the plateau depends on the corrosion rate and therefore does not permit to measure of the amount of oxide present.

© 2007, first edition, EPFL Press

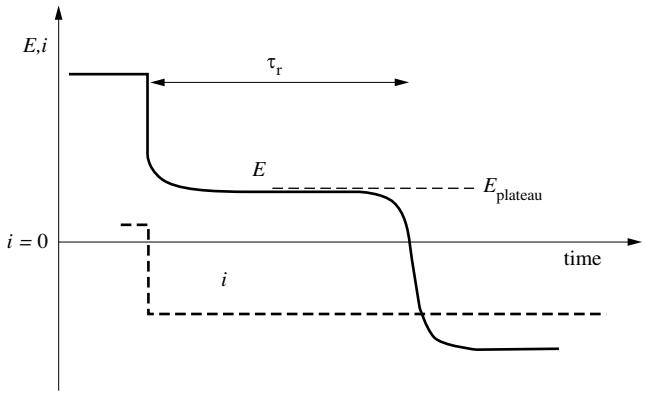


Figure 6.18 Potential transient for passive film reduction at constant current. The reduction time τ_r is proportional to the reduction charge of the film.

Impedance spectroscopy is a widely used electrochemical in situ method for the characterization of passive films. It yields information on the electrical properties, for example the flat band potential and the charge carrier density in semiconducting oxides (Section 3.4.4). The technique can also be used for monitoring changes in the film thickness. The interpretation of impedance data measured on passive films is often delicate and therefore can not be discussed in more detail here. In recent years important structural information on passive films has come from the in situ application of ***scanning probe techniques***, notably the STM (Section 3.4.2). These methods are able to give structural information on an atomic scale, and they can be used to monitor slow structural transformations of passive films with time. Another in situ method, the ***electrochemical quartz crystal microbalance*** (EQCM), can be used for monitoring tiny mass changes related to film growth and dissolution. For this one observes the shift in resonance frequency of a quartz crystal on which a thin metal film has been deposited that acts as working electrode.

Among the *ex-situ techniques,* surface analysis methods, in particular XPS, AES and SIMS, have been widely used for characterizing the chemical composition of passive films. When these techniques are employed, it is usually assumed that the film undergoes no significant chemical modification during the transfer from the electrolyte to the high-vacuum analysis chamber, an assumption that needs to be critically examined in each case.

The application of ex situ techniques for the study of the structure of passive films suffers from the fact that the films, which are only a few atom layers thick, may undergo structural changes when removed from the electrolyte or when submitted to intense radiation. The effect of radiation can be a problem in structural studies using electron diffraction in a transmission electron microscope. In ***EXAFS*** (Extended X-ray Absorption Fine Structure), another technique for the study of passive oxide films, one uses synchrotron radiation. EXAFS allows one, in principle, to determine the atomic number, distance and coordination number of the atoms surrounding the element whose absorption edge is being examined. Scanning probe techniques such

© 2007, first edition, EPFL Press

as STM and AFM have been applied *ex situ* as well as *in situ*. They are the preferred methods today for obtaining structural information. For example, passive films on nickel were found to have a crystalline structure with lattice parameters corresponding to NiO [12]. For a more detailed discussion of the different methods the reader is referred to the general references listed at the end of this chapter.

Structure of passive films

The atomic structure and the stoichiometry of passive films on a given metal vary with their formation conditions. Some passive films that form on corrosion-resistant metals have an amorphous structure. The structure and growth rate of an amorphous film does not reflect the grain orientation of the substrate and, as a result, we expect amorphous films to be more homogeneous than crystalline films. This should be beneficial for their ability to protect the metal. Using electron diffraction it has been found that the passive films on iron in sulfuric acid are crystalline, whereas those on iron-chromium alloys with high chromium content are amorphous [13]. The crystalline nature of passive films on iron and nickel has also been confirmed by STM. Figure 6.19 shows an ex situ STM image of the passive oxide film on iron formed in a borate buffer solution on a polycrystalline sputter deposited metal film [14]. The regular arrangement and spacing of the oxygen ions indicates a [111] orientation of the passive film. Similar images were observed by the same authors by in situ observation. These results clearly demonstrate that passive films on iron formed at neutral pH are crystalline. More recent studies using the STM indicate that the degree of crystallinity of passive films may vary with time and with potential. For example, a freshly formed passive film on a (110) oriented chromium single crystal at a low passive potential

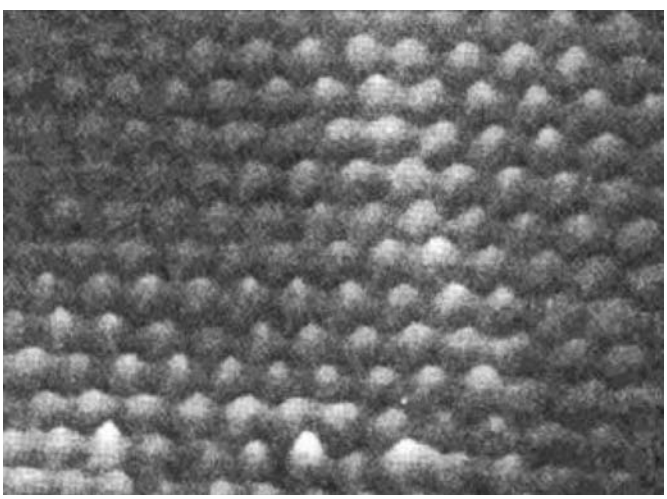


Figure 6.19 Ex situ STM image with atomic resolution of the surface of passive film formed on iron by anodic polarization in a borate buffer solution of pH 8.4. The image demonstrates the crystalline nature of the oxide film [14].

© 2007, first edition, EPFL Press

was found to be amorphous and consisting mostly of hydroxide. [15]. With increasing time the film became increasingly crystalline, oxide growing underneath the hydroxide. The transformation occurred the faster the higher the applied potential. A hydroxide layer remained at the outer surface at all times. Structural transformations with time and potential have been observed on other passive films as well and seem to be a general characteristic of passive films. Apparently, there is no simple relationship between the crystallinity of passive films and their ability to confer corrosion resistance to the underlying metal.

Chemical composition of passive films

Passive films are formed of metal oxides the composition of which can vary with experimental conditions. Typically, hydroxyl groups are found at the outer surface of the film. Their presence is due to the fact that the oxide surface is hydrated, or it may result from the precipitation of a thin hydroxide layer from the solution. Often during growth, the passive film traps small amounts of anions from the electrolyte: Cl^-, SO_4^{2-}, PO_4^{3-}, etc. These anions can occupy the lattice sites of O^{2-} ions or they are located at defects. If it is easy to imagine a chloride ion occupying the lattice site of an O^{2-} ion, it is more difficult to visualize the much larger sulfate ion occupying such a position. XPS spectra of passive films sometimes show the presence of bound water, which manifests itself as a broadening of the oxygen peak. This indicates that the oxides forming the passive film are hydrated.

Surface analysis methods are particularly useful for the study of the concentration and of the oxidation state of cations present in passive films. Most often, the cation oxidation state in passive films corresponds to that of the most stable oxide: Ni^{2+}, Al^{3+}, Fe^{3+}, Cr^{3+}, Ta^{5+}. Sometimes, the oxidation state varies with the potential. For example, a passive film of nickel formed in alkaline solution may contain Ni^{3+} cations at high potentials. A film of Cr_2O_3 transforms into CrO_3 at the potential E_b. The high solubility of the hexavalent oxide causes the dissolution of the passive film (Section 6.3). Sometimes, the oxidation state of the cations in a passive film varies as a function of distance from the metal surface. Passive films of iron, for example, are crystalline and comprised mainly of Fe_2O_3. They contain, however, a certain proportion of Fe^{2+} ions near the metal-oxide interface. In fact, certain authors have attributed a duplex structure to the passive film on iron: an outer layer of γ-Fe_2O_3 and an inner layer of Fe_3O_4. Magnetite, Fe_3O_4, contains both divalent and trivalent iron cations, its composition corresponding to the formula Fe_2O_3-FeO. Given that the passive film on iron is only a few atomic layers thick, rather than consisting of two well-defined oxides, it is more probable that its stoichiometry varies in a continuous fashion. The stoichiometry of the film is then described by the formula $Fe_{3-y}O_4$, where y increases from zero to one third as one moves from the metal-film to the film-solution interface [16].

The chemical composition of passive films formed on alloys is even more complicated because two or more metal cations are present in various amounts. Indeed, the concentration ratio of cations in passive films on alloys often differs

© 2007, first edition, EPFL Press

substantially from that expected from the alloy composition. Two phenomena, selective oxidation of alloy elements and selective dissolution of film constituents can contribute to this. The behavior will be illustrated for passive films on iron-chromium alloys.

Passive films on iron-chromium alloys

Passive films formed on iron-chromium alloys in acid solution exhibit generally a higher relative chromium content than the underlying alloy. This is evidenced by the XPS data of Figure 6.20 showing the atomic ratio Cr / (Fe+Cr) in the passive film as a function of the same ratio in the alloy for passive films formed in sulfuric acid [17,18]. Over a narrow concentration range above about 12% the chromium concentration in the film increases dramatically. This leads to a marked increase in the corrosion resistance of the alloys. Indeed, one of the most common stainless steels, AISI 430, has the composition Fe-17Cr, well in the region of strong chromium enrichment of the passive film.

The dashed line in Figure 6.20 shows a theoretical prediction from a simple model applied to a binary alloy AB [19]. This model takes into account the fluxes at the metal-film and film-electrode interfaces. The flux of element k at the metal-film interface, $N_{k,m}$, is equal to:

$$N_{k,\,m} = k_{k,\,m(E)}\,\rho_m\,X_{k,\,m} \tag{6.10}$$

where $X_{k,m}$ designates the mole fraction of the element k in the alloy ($k = A,B$); ρ_m is the density of the alloy; and $k_{k(E)}$ is a rate constant whose value depends on the

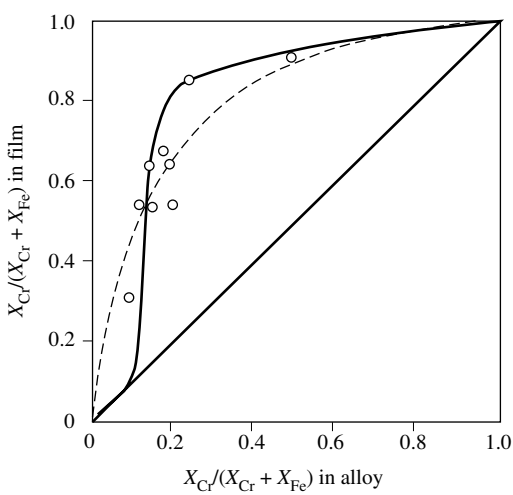


Figure 6.20 Average chromium content of passive film of Fe-Cr alloys, measured with XPS, as a function of Cr concentration in the alloy. Solution is 0.5M H_2SO_4. The dashed curve is calculated using equation (6.13) [13,15].

© 2007, first edition, EPFL Press

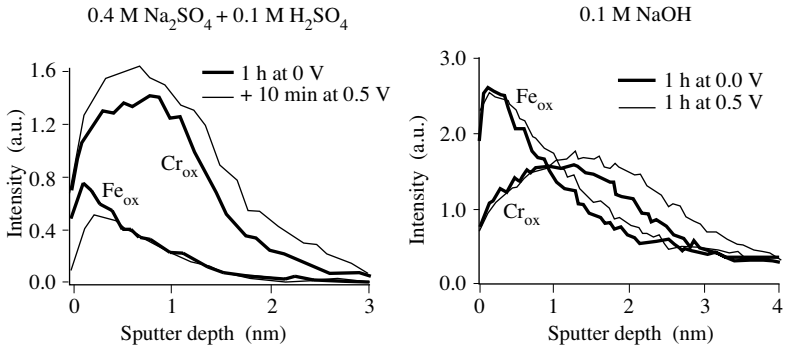


Figure 6.21 Auger depth profiles measured on Fe-25Cr alloy anodically passivated at 0 V and at 0.5 V in solutions of 0.4 M Na_2SO_4 + 0.1 M H_2SO_4 and of 1 M NaOH, respectively. In acid solution the passive film is enriched in chromium oxide while in alkaline solution it contains mostly iron oxide [21].

potential. In the same fashion, the flux of the element k at the film-electrolyte interface, $N_{k,f}$, is equal to:

$$N_{k,\,f} = k_{k,\,f(E)}\,\rho_{ox}\,X_{k,\,f} \tag{6.11}$$

At steady state, $N_{k,m} = N_{k,f}$, which gives:

$$\frac{k_{A,\,m(E)}}{k_{B,\,m(E)}}\,\frac{X_{A,\,m}}{X_{B,\,m}} = \frac{k_{A,\,f(E)}}{k_{B,\,f(E)}}\,\frac{X_{A,\,f}}{X_{B,\,f}} \tag{6.12}$$

and

$$\frac{X_{A,\,f}}{X_{B,\,f}} = r\,\frac{X_{A,\,m}}{X_{B,\,m}} \tag{6.13}$$

$$r = \frac{k_{A,\,m(E)}}{k_{B,\,m(E)}}\,\frac{k_{B,\,f(E)}}{k_{A,\,f(E)}}$$

The equation (6.13) is plotted in Figure 6.20 with the (arbitrary) assumption r = 8. The model permits to rationalize the observed chromium enrichment in terms of the relative rates of oxidation and dissolution reactions of alloy components, but the agreement with experimental data is mediocre. One reason is that the model ignores the fact that the passive film composition on Fe-Cr alloys is not uniform and stationary, but varies as a function of depth and time [20].

The chemical composition of passive films on Fe-Cr alloys depends strongly on the corrosive environment and on the applied potential. In acid media the films become enriched in chromium oxide, while in alkaline media they contain more iron oxide. The Auger depth profiles of Figure 6.21 illustrate this observation. They were measured on a Fe-25Cr alloy after 1 hour anodic polarization in an acid sulfate

© 2007, first edition, EPFL Press

solution and in 0.1M NaOH respectively. One also observes that the composition varies with depth and depends on applied potential. The difference in film composition between acid and alkaline media results from the fact that in acids chromium oxide is almost insoluble, contrary to iron oxide. Upon anodic polarization of a Fe-Cr alloy iron is therefore preferentially dissolved from the growing film leading to chromium enrichment. This behavior has been confirmed by electrochemical quartz crystal microbalance (EQCM) experiments; during anodic film growth on Fe-Cr in sulfuric acid the total mass of the electrode decreases rather than increases because of selective dissolution of iron [21]. In alkaline solution this mechanism does not take place, because iron oxide is only sparingly soluble. As a consequence, the passive films formed on Fe-Cr alloys in alkaline media contain predominantly iron oxide.

It follows from these and many other data that passive films formed on a given metal or alloy do not possess a unique structure, thickness and stoichiometry. Rather, these parameters depend on environmental conditions such as pH and chemical composition of the electrolyte, applied potential and exposure time. Passive films should be seen as dynamic entities, capable of repairing themselves when damaged, and whose structure and composition depend on prevailing environmental conditions as well as on the base material.

6.2.2 Growth of anodic oxide films

The growth and dissolution of passive oxide films involve ion-transfer reactions at the metal-film and the film-electrolyte interfaces (Figure 6.22). At the metal-film interface, metal ions are produced that occupy the cationic sites in the oxide.

$$M \rightarrow M^{n+}_{(ox)} + n\,e \tag{6.14}$$

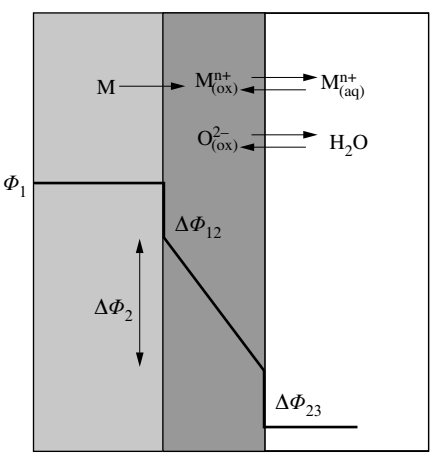


Figure 6.22 Variation of potential across a passive film (schematic).

© 2007, first edition, EPFL Press

Under the influence of the strong electric field present in the film (typically 10^8 V m^{-1}), the cations migrate toward the outer surface, while the anions migrate in the opposite direction. At the film-electrode interface, two charge-transfer reactions are possible: the movement of metal ions into solution (6.15) and the deposition (or dissolution) of O^{2-} ions (6.16).

$$M^{n+}_{(ox)} \rightarrow M^{n+}_{(aq)} \tag{6.15}$$

$$O^{2-}_{(ox)} + 2\,H^{+}_{(aq)} \leftrightarrows H_2O \tag{6.16}$$

Reaction (6.16) controls whether the film grows or thins.

The rate of the reactions (6.15) and (6.16) depends on the potential difference $\Delta\Phi_{23}$ at the film-electrolyte interface. The ionic current density corresponding to metal-ion dissolution i_{diss} is thus described by the equation (6.17) and the current density for film growth (or thinning), i_{gr}, by (6.18).

$$i_{diss} = n\,F\,k_d \exp(B_d\ \Delta\Phi_{23}) \tag{6.17}$$

$$i_{gr} = n\,F\,k_{gr} \exp(B_{gr}\ \Delta\Phi_{23}) - n\,F\,k'_{gr} \exp(-B'_{gr}\ \Delta\Phi_{23}) \tag{6.18}$$

Here, the terms k_d, k_{gr} et k'_{gr} , $B_{d,}$ B_{gr} et B'_{gr} are constants. The expressions (6.17) and (6.18) resemble the Butler-Volmer equation for charge transfer at the metal-electrolyte interface, keeping in mind that in equation (6.17) the cathodic term can be neglected. The potential difference $\Delta\Phi_{23}$ is not directly measurable. It depends, among other, on the pH of the solution.

High field conduction

The ionic current density in the oxide is an exponential function of the electric field. The model shown in Figure 6.23 represents schematically the effect of electric field on the activation energy that charge carriers must overcome in order to move through the film. Let us assume that the ionic current is only due to cation migration from the metal-film interface towards the film exterior. The detailed mechanism of the movement (through interstitials or vacancies) is unimportant; it suffices to consider a cation as it moves from one position of minimum energy to the next. The electric field reduces the activation energy of the jump, $\Delta G^{\#}$, in the direction of the field.

$$\Delta G^{\#} = \Delta G^{\#}_{ch} - \frac{\alpha' z_+ F\,a\Delta\Phi_2}{L} \tag{6.19}$$

In the opposite direction, the field increases the activation energy

$$\Delta G^{\#}_{inv} = \Delta G^{\#}_{ch} + \frac{(1-\alpha')\,z_+ F\,a\Delta\Phi_2}{L} \tag{6.20}$$

Here, L refers to the thickness of the film and $\Delta\Phi_2$ is the potential difference across the film. The quantity $\Delta G^{\#}_{ch}$ represents the activation energy in the absence of an

© 2007, first edition, EPFL Press

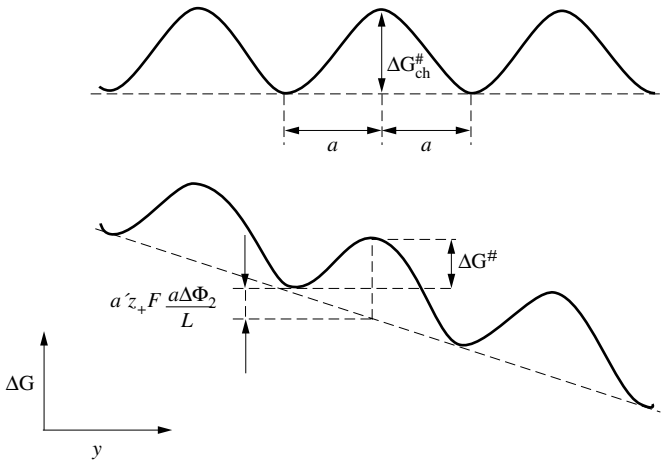


Figure 6.23 Effect of electrical potential on activation energy of mobile ions in oxide film.

electric field; α' is a proportionality constant whose value varies with the symmetry of the energy barrier ($0 < \alpha' < 1$); z_+ designates the charge on the ion ($z_+ = n$ for the cations) and a represents half the distance between two equilibrium positions of the ion.

The effect of the electric field allows us to neglect cation movement in the reverse direction. The flux of cations in the oxide, N_+, is then given by (6.21), where ν corresponds to a frequency factor. The quantity c_+ represents the concentration of mobile charge carriers, typically point defects, assumed constant within the film.

$$N_+ = 2\,a\nu\,c_+ \exp\left(-\frac{\Delta G^{\#}}{RT}\right) \tag{6.21}$$

With (6.19) we obtain (6.22).

$$N_+ = 2\,a\nu c_+ \exp\left(-\frac{\Delta G^{\#}}{RT}\right)\exp\left(\frac{\alpha' z_+ F a}{RT}\,\frac{\Delta\Phi_2}{L}\right) \tag{6.22}$$

The current density in the film corresponds to:

$$i = n F N^+ \tag{6.23}$$

By grouping together the quantities that are independent of the electric field into the constants i_0' and B', we obtain the ***high field conduction equation***:

$$i = i_0' \exp\left(B'\frac{\Delta\Phi_2}{L}\right) \tag{6.24}$$

© 2007, first edition, EPFL Press

with

$$i_0' = n F 2 a v c_+ \exp\left(-\frac{\Delta G_{\mathrm{ch}}^{\#}}{RT}\right)$$

$$B' = \frac{\alpha' z_+ F a}{RT}$$

The ratio $\Delta\Phi_2/L$ represents the average electric field in the film. Often, in order to simplify equation (6.24) the potential differences at the metal-film and film-solution interfaces are neglected, and one sets: $\Delta\Phi_2 = \Delta\Phi$.

According to (6.24), the ionic current density in the film varies exponentially with the electric field. Even though this relation has been derived here from a rather simple model, it holds true quite generally. We therefore can also look at equation (6.24) as an empirical equation that describes the relation between the ionic current, the potential and the thickness of solid oxide films, in a similar way as the Butler-Volmer equation describes the relation between current and potential for a metal-electrolyte interface.

Galvanostatic growth of oxide films

To achieve anodic film growth, two methods are widely used: (i) application of a constant current (galvanostatic oxidation), (ii) application of an anodic potential (potentiostatic oxidation).

According to Faraday's law, if dissolution is negligible, the rate of film growth is proportional to the applied current density.

$$\frac{\mathrm{d}L}{\mathrm{d}t} = k_{\mathrm{F}} i \tag{6.25}$$

$$k_{\mathrm{F}} = \frac{M_{\mathrm{ox}}}{n F \rho_{\mathrm{ox}}}$$

Here, k_{F} is a parameter that depends on the molecular mass of the film, M_{ox}, and on its density, ρ_{ox}.

During galvanostatic oxidation, the current density i remains constant. Differentiating equation (6.24) therefore gives:

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \frac{B'}{\ln \frac{i}{i_0'}} \frac{\mathrm{d}\Delta\Phi_2}{\mathrm{d}t} \tag{6.26}$$

By combining equations (6.26) and (6.25) we can eliminate $\mathrm{d}L/\mathrm{d}t$. After integration, we obtain:

$$\Delta\Phi_2(t) = k_{\mathrm{i}} t \tag{6.27}$$

© 2007, first edition, EPFL Press

where

$$k_i = \frac{k_F i}{B'} \ln \frac{i}{i'_0}$$

The electrode potential measured as a function of time, $E(t)$, corresponds to $\Delta\Phi_2(t)$, to within a constant, provided that the potential differences at the film-metal and film-electrolyte interfaces do not vary.

$$E(t) = \Delta\Phi_2(t) + \text{constant} \tag{6.28}$$

According to (6.27) and (6.28), the measured potential increases linearly with the time of oxidation and so does the film thickness, which is proportional to $\Delta\Phi_2$.

$$L(t) = k_G \, \Delta\Phi_2(t) \tag{6.29}$$

According to (6.26) the proportionality constant k_G is equal to:

$$k_G = \frac{B'}{\ln \frac{i}{i'_0}} \tag{6.30}$$

Figure 6.24 shows the results of a galvanostatic oxidation of tantalum [22].This metal forms insulating oxide layers whose thickness may reach several hundred nanometers, which corresponds to potentials greater than 100 V. The anodization of tantalum is used in numerous applications, including the fabrication of electrolytic capacitors. Aluminum, titanium and zirconium show similar behavior, even if the voltages never reach the levels shown in Figure 6.24. Metals that form weakly

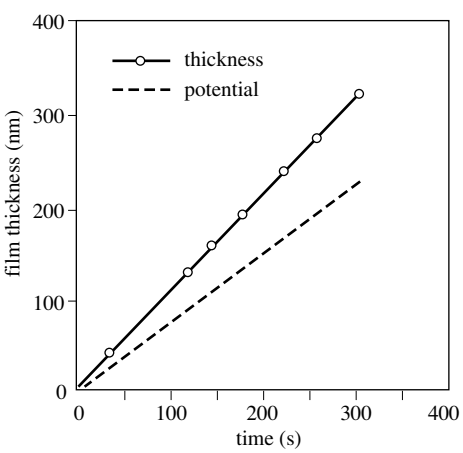


Figure 6.24 Galvanostatic oxide film growth on tantalum: variation of thickness and of potential as a function of the anodization time. Electrolyte: 1% Na_2SO_4, T = 19 °C, applied current density is 8 mA cm^{-2} [22].

© 2007, first edition, EPFL Press

conducting anodic oxide films that support anodization to many volts are referred to as ***valve metals*** in the electrochemical literature.

Potentiostatic growth of oxide films

At constant potential and as long as the film does not dissolve, the film growth rate is described by equation (6.31):

$$\frac{\mathrm{d}L}{\mathrm{d}t} = k_{\mathrm{F}}\, i = k_{\mathrm{F}}\, i_0' \exp\left(B' \frac{\Delta\Phi_2}{L}\right) \tag{6.31}$$

This equation can only be integrated by numerical methods. The expression (6.32) gives an approximate analytical solution, valid for $B'\Delta\Phi_2 >> L$ [23]:

$$\frac{1}{L} = A_1 - A_2 \ln t \tag{6.32}$$

A_1 and A_2 are constants. The current density therefore varies with time according to the following relation:

$$\ln i = A_1' - A_2' \ln t \tag{6.33}$$

The values of the constants A_1' and A_2' depend on the applied potential. Figure 6.25 shows the experimental results of a potentiostatic oxidation of tantalum. The approximate solution given by (6.32) is also presented.

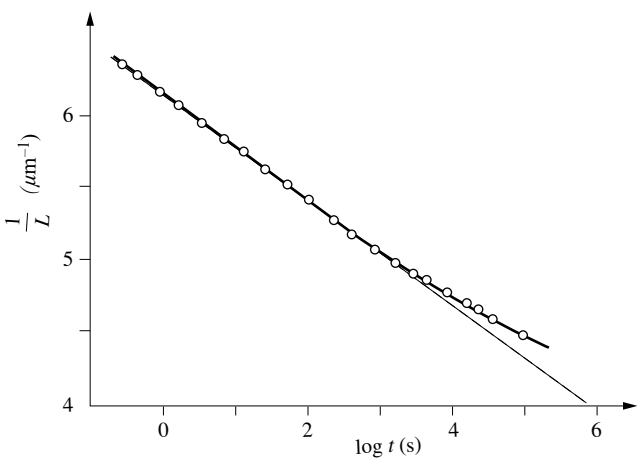


Figure 6.25 Potentiostatic oxide film growth on Tantalum: variation of $1/L$ (L = thickness) with the anodization time. Electrolyte : 1% Na_2SO_4, T = 19 °C, applied potential is 120 V. The dashed line corresponds to equation (6.32) [22].

© 2007, first edition, EPFL Press

When there is no film dissolution, the current keeps decreasing and a true steady state is never established. After a sufficiently long period of time, however, the anodic current density reaches such small values that for all practical purposes the change in film thickness becomes negligible.

Influence of film dissolution

The total current density i at the film-electrolyte interface is the sum of the growth current density i_{gr} (6.18) and the dissolution current density i_{diss} (6.17).

$$i = i_{gr} + i_{diss} \tag{6.34}$$

The change in film thickness in presence of film dissolution is given by

$$\frac{dL}{dt} = k_F\,(i - i_{diss}) \tag{6.35}$$

Under steady state conditions the growth current is nil and therefore $i_{gr} \approx i_{diss}$. With (6.35) and (6.24), we obtain for this situation:

$$i \approx i_{diss} = i_0' \exp\left(\frac{B'\,\Delta\Phi_2}{L}\right) \tag{6.36}$$

In the passive region, the dissolution current density i_{diss} in general varies little with the applied potential. Rearranging (6.36) and taking the logarithm, we obtain:

$$L = k_p\,\Delta\Phi_2 = k_p E + \text{constant} \tag{6.37}$$

with

$$k_p = \frac{B'}{\ln \frac{i_{diss}}{i_0'}}$$

Here, we used the relation $\Delta\Phi_2 = E$ + constant, where E is the electrode potential. According to equation (6.37), under steady state dissolution conditions, the thickness of the passive film increases linearly with the applied potential.

A linear relation between film thickness and the applied potential is also found in absence of dissolution, provided that the polarization time is long enough for the rate of film growth to drop to close to zero. Figure 6.26 shows results obtained by *XPS* on iron and Fe-Cr alloys in a sodium hydroxide solution. The films do not dissolve in this environment, and a true steady state is therefore not established. Nevertheless, good linearity is observed between the film thickness and applied potential after a period of five minutes already.

© 2007, first edition, EPFL Press

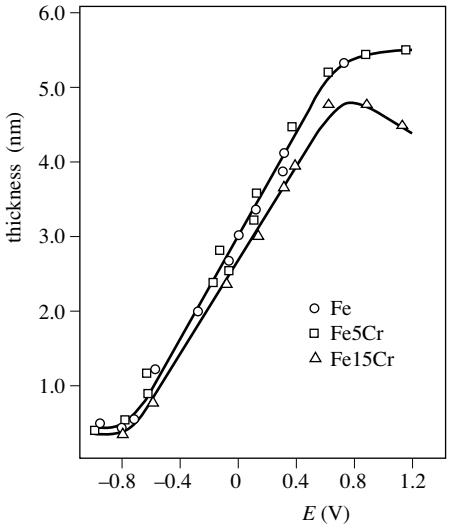


Figure 6.26 XPS film thickness on Fe and Fe-Cr alloys at different potentials [24].

Interface controlled film growth

According to the high filed conduction model the growth rate of oxide films is controlled by the migration of ions in the solid under the effect of a strong electric field. When we apply a potential step, the field strength $\Delta\Phi_2/L$ diminishes as the film thickens and as a consequence, the growth rate decreases. The potential differences at the metal-film and film-electrolyte interfaces, $\Delta\Phi_{12}$ and $\Delta\Phi_{23}$, do not significantly change during growth. A different model can be proposed, where the growth rate is controlled by the rate of the charge transfer reaction at the oxide-solution interface [25,26]. In this case, the electric field in the oxide stays constant during film growth at constant potential, while the potential difference at the controlling interface, $\Delta\Phi_{23}$, diminishes with increasing film thickness. Figure 6.27 schematically shows the potential profiles predicted by the two models when the potential of a passive electrode is stepped anodically. It is assumed here that the charge transfer at the film-electrolyte interface is controlling, but the same reasoning can be applied to a situation where the reaction at the metal-film interface is rate controlling. Models combining both types of rate control have also been proposed in the literature [27]. The intermediate states for the two models shown in Figure 6.27 differ, but the final states are the same, depending only on the applied potential. The mathematical development of the interface control model shows that if the interfacial reaction obeys Butler-Volmer kinetics, the predicted growth curve is very similar to that of the high field conduction model. Recent experiments with the electrochemical quartz crystal microbalance indicate that for thin passive films on chromium the film growth rate is predominantly controlled by an interfacial reaction [28], while film growth on valve metals, which leads to thicker films, is governed by high field conduction [29].

© 2007, first edition, EPFL Press

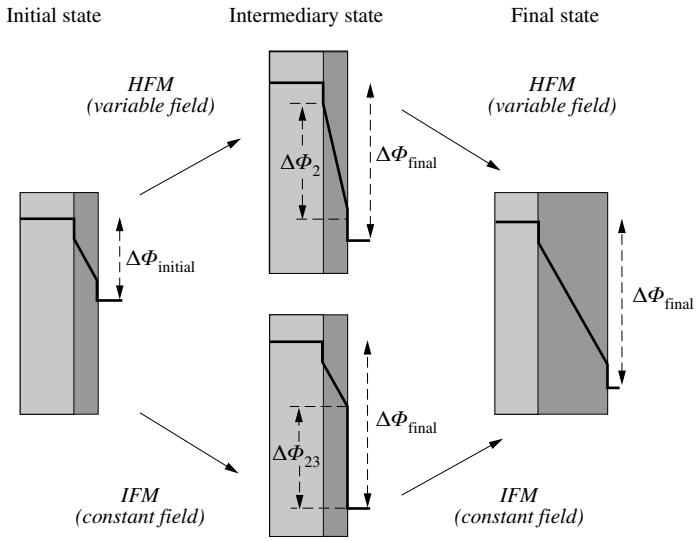


Figure 6.27 Evolution of the potential profile during potentiostatic film growth controlled, respectively, by high field conduction (HFM) and by charge transfer at the film-electrolyte interface (IFM). When an anodic potential step is applied to a passive metal electrode, both mechanisms predict a decreasing growth rate with time.

6.2.3 Electron transfer reactions

Electrode reactions taking place on passive metals can be divided into two categories: ***ion-transfer reactions***, responsible for film growth and dissolution; and ***electron-transfer reactions***, which involve the oxidation or reduction of a component of the electrolyte.

The electron-transfer reaction (6.38) is of particular importance in corrosion and electrochemical technology. It can take place on active as well as passive metal surfaces.

$$O_2 + 4\,H^+ + 4\,e = 2\,H_2O \tag{6.38}$$

For this or other electron transfer reactions to take place on a passive metal, the oxide film present must exhibit electronic conductivity. Depending on the prevailing potential with respect to the reversible potential of the oxygen electrode, reduction of oxygen or oxidation of water can then be achieved.

Oxidation of water

The thickness of passive films normally increases with the applied potential. For chemically stable, non-conducting films, it can reach several tens of nanometers at anodization potentials of 100 V or more. On the other hand, on films exhibiting

© 2007, first edition, EPFL Press

electronic conductivity, electron-transfer reactions can take place at the film-electrolyte interface and therefore above a certain potential, water is oxidized into oxygen. A mainly electronic current then crosses the interface and film growth ceases. The potential required for oxidation of water generally is on the order of 1.2 to 2 volts and the thickness of electron conducting films therefore does not exceed a few nm. The passive film formed on nickel in sulfuric acid is a typical example of this type of behavior.

Reduction of oxygen

When a corrosion cell is formed between an active anode and a passive cathode (see Chap. 7), oxygen reduction takes place on the passive film. The rate of oxygen reduction depends on the electronic properties of the film in a complicated way, the reason being that the oxygen reduction mechanism involves several steps. Figure 6.28 [30] illustrates the fact that the reduction rate on a passive oxide films depends not only on the film properties, but also on the nature of the oxidizing agent.

Cathodic partial current densities are shown corresponding to the reduction of O_2, Fe^{3+} and Ce^{4+} ions, respectively, together with the anodic partial current density of austenitic stainless steel that serves as electrode. The electrolyte is sulfuric acid at 60 °C. The concentration of the various oxidizing agents is chosen such that their limiting current densities have comparable values, on the order of 100 μA cm^{-2}. In Figure 6.28, this corresponds to the plateaus observed near the passivation potential. In the passive potential region, the cathodic partial current density for the reduction of oxygen is clearly lower than that at the passivation potential. In other words, the reduction of oxygen slows significantly as the stainless steel passes from the active to

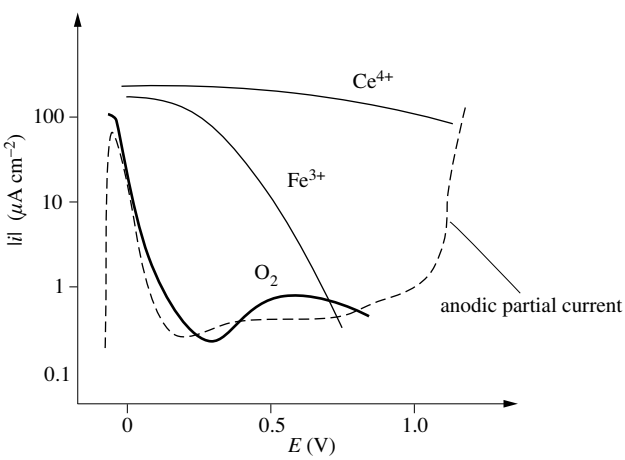


Figure 6.28 Cathodic partial current density for the reduction of oxygen, of Fe^{3+} ions, and of Ce^{4+} ions on an austenitic stainless steel of type AISI 304 in 5% H_2SO_4. Concentration of the oxidizing agents: (3/*n*) mM l^{-1} (*n*= charge number). Dashed line: anodic partial current density [30].

© 2007, first edition, EPFL Press

the passive state. In contrast, the partial current density for Ce^{4+} reduction is not significantly influenced, its value in the passive region being close to the limiting current density. The reduction of Ce^{4+} to Ce^{3+} involves the transfer of a single electron and therefore is a much simpler reaction than the reduction of O_2. The reduction of Fe^{3+} to Fe^{2+} also involves a single electron transfer and the cathodic partial current density in the low passive region exhibits a plateau, similar as for Ce^{4+} reduction. However, at higher potentials in the passive region the partial current density drops. The reason is that the standard potential of the Fe^{3+}/Fe^{2+} couple ($E° = 0.77$ V) is located well inside the passive region, and is much lower than that of Ce^{4+}/Ce^{3+} ($E^o = 1.72$V).

On passive metals, the reduction kinetics of oxidizing agents therefore depends on two factors: the electronic properties of the oxide, and the chemical nature of the oxidizing agents.

Electron transfer by tunneling

When the passive film present on a metal is sufficiently thin (≤ 2-3 nm), electrons exchanged between the metal and the electrolyte can cross the film by the ***tunneling***. In this case, the reaction rate decreases as the film becomes thicker. On the other hand, the exchange rate does not depend, in principle, on the electronic properties of the film. Figure 6.29 shows polarization curves for the oxidation of ferrocyanide ions and for the reduction of ferricyanide ions measured on passive iron electrodes with different film thicknesses [31].

$$Fe(CN)_6^{3-} + e = Fe(CN)_6^{4-} \tag{6.39}$$

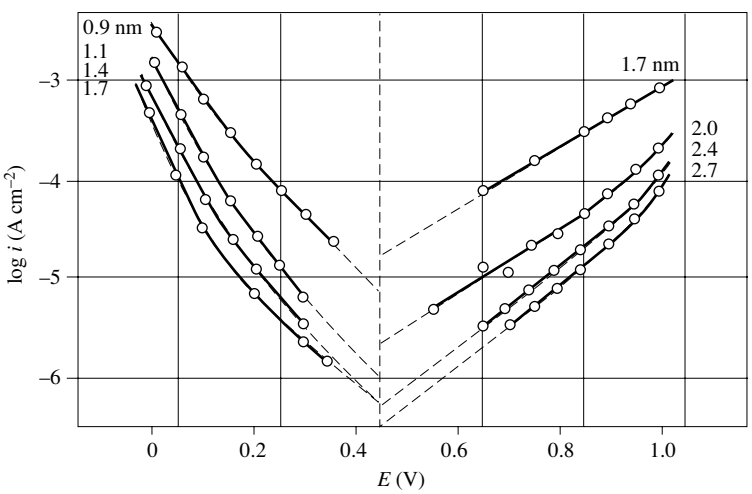


Figure 6.29 Effect of passive film thickness on kinetics of electron tansfer between $Fe(CN)_6^{3-}$ and $Fe(CN)_6^{4-}$ on passive iron at pH 8.4 [31].

© 2007, first edition, EPFL Press

The exchange current density (given by the intersection between the anodic and cathodic Tafel lines) decreases as the thickness of the passive film changes from 1.7 to 2.7 nm. The behavior can be explained by the reduced probability of electron tunneling through the oxide barrier with increasing film thickness.

Electronic properties of oxides

When an oxide film is too thick for electron tunneling to take place, the intrinsic electronic properties of the film control the kinetics of the charge transfer at the film-electrolyte interface. Most passive films are semiconducting, but their electronic properties can vary greatly from one to another. Figure 6.30 gives the energies of the valence and conduction bands for different oxides, relative to the vacuum level. It also indicates the standard potentials of several redox reactions for the species present in the electrolytes, on the electrochemical scale.

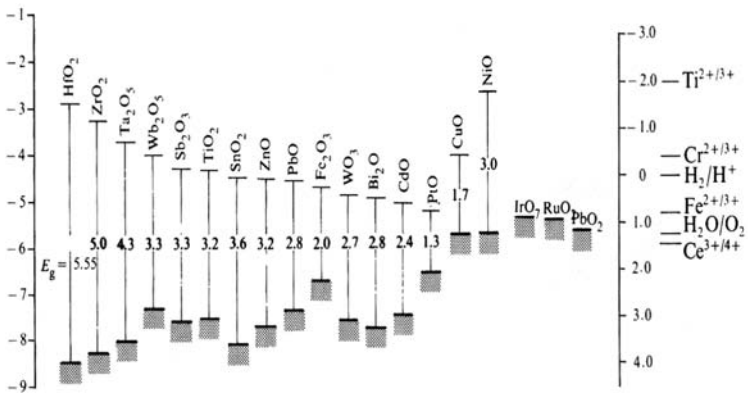


Figure 6.30 Band energy of different semiconducting oxides and standard potentials of electron-transfer redox reactions [31].

The oxides IrO_2, RuO_2 and PbO_2 are metallic conductors. They therefore do not have a forbidden band. The oxides of Ti, Fe, Sn, Sb, Nb, Bi, W are *n*-type semiconductors, while the oxides of Cu, Ni, and Cr are *p*-type semiconductors. Their forbidden band is on the order of 2 to 3 eV. The oxides of Ta, Zr, Hf and Al are insulators because the forbidden band energy exceeds 4 eV.

Electron transfer between a semiconducting oxide and the electrolyte

Electron transfer reactions at interfaces between a semiconducting oxide and an electrolyte can take place in two ways:

- electron exchange with the conduction band (n-type semiconductors);
- electron exchange with the valence band (p-type semiconductors).

To study these mechanisms we consider the equilibrium between ferrous and ferric ions at a semiconductor electrode. This redox system involves a single electron transfer.

© 2007, first edition, EPFL Press

$$Fe^{3+} + e = Fe^{2+} \text{ (conduction band)} \tag{6.40}$$
$$Fe^{3+} = Fe^{2+} + h \text{ (valence band)} \tag{6.41}$$

For a *n-type semiconductor* the anodic partial current density, $i_{a,cond}$, corresponds to:

$$i_{a,\,cond} = F\,k_{a,\,cond}\,N_{cond}\,c_{Fe}^{2+}\,D_{red} \tag{6.42}$$

The subscript "*cond*" means "conduction band". According to (6.42) $i_{a,cond}$ is proportional to the concentration of unoccupied electronic states in the conduction band, N_{cond}, and to the concentration of occupied states in the electrolyte: c_{Fe2+} D_{red}. Here, c_{Fe2+} is the concentration of Fe^{2+} ions in solution and D_{red} represents the density of states function.

In the expression (6.42), there are no parameters that depend on the applied potential, thus demonstrating the following equality:

$$i_{a,\,cond} = i_{a,\,cond}^{0} \tag{6.43}$$

where $i_{a,cond}^{0}$ depends only on system properties.

For an *n*-type semiconductor, the cathodic partial current density, $i_{c,cond}$, is proportional to the concentration of electrons in the conduction band at the surface of the semiconductor, $c_{n,s}$, and to the concentration of Fe^{3+} ions in solution, c_{Fe3+}.

$$i_{c,\,cond} = -F\,k_{c,\,cond}\,c_{n,\,s}\,c_{Fe3+}\,D_{ox} \tag{6.44}$$

In this equation, D_{ox} represents the density of states function, and the product $c_{Fe3+}\,D_{ox}$ corresponds to the concentration of unoccupied electronic states in solution.

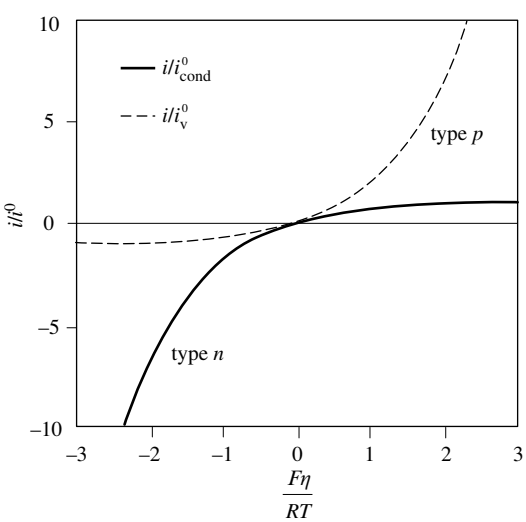


Figure 6.31 Dimensionless polarization curves corresponding to an electron transfer reactionwith the conduction band (*n*-type semiconductor) or to a hole exchange reaction with the valence band (*p*-type semiconductor).

© 2007, first edition, EPFL Press

The electron concentration at the semiconductor surface, $c_{n,s}$, depends on the potential difference across the space charge layer, $\Delta\Phi_{SC}$.

$$c_{n,s} = c_{n,b} \exp\left(-F \frac{\Delta\Phi_{SC}}{RT}\right) \tag{6.45}$$

$$\Delta\Phi_{SC} = \Phi_{SC,b} - \Phi_{SC,s}$$

Here, $c_{n,b}$ represents the electron concentration; $\Phi_{SC,b}$ and $\Phi_{SC,s}$ are the potentials in the bulk and at the surface of the semiconductor, respectively (Sect. 3.5).

The polarization of a semiconducting electrode most of all modifies $\Delta\Phi_{SC}$. This leads to a relation for the overpotential:

$$\eta \approx \Delta\Phi_{SC} - \Delta\Phi^{\circ}_{SC} \tag{6.46}$$

where $\Delta\Phi^{\circ}_{SC}$ refers to the potential difference across the space charge layer at equilibrium. With the definition,

$$i^{\circ}_{c,\,cond} = F\, k_{c,\,cond}\, c_{n,\,b}\, c_{Fe^{3+}}\, D_{ox} \exp\left(-F \frac{\Delta\Phi^{\circ}_{SC}}{RT}\right) \tag{6.47}$$

equation (6.44) becomes:

$$i_{c,\,cond} = -i^{\circ}_{c,\,cond} \exp\left(-\frac{F\eta}{RT}\right) \tag{6.48}$$

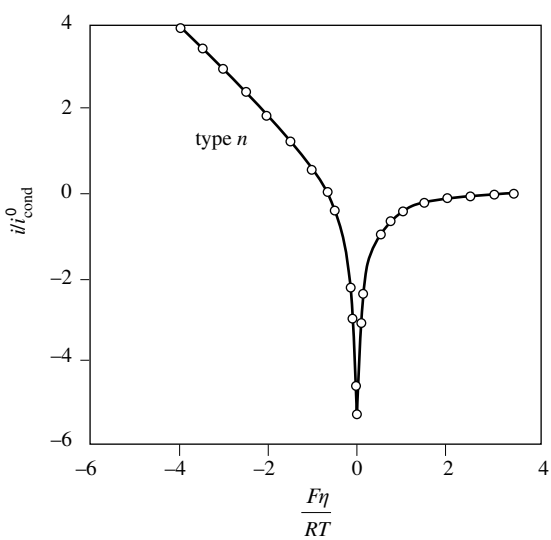


Figure 6.32 Nondimensional calculated polarization curve, presented on a logarithmic scale, for redox reaction at n-type semiconductor electrode.

© 2007, first edition, EPFL Press

The measured current density, i, is equal to the sum of the anodic and cathodic partial current densities:

$$i = i_{\text{a, cond}} + i_{\text{c, cond}} \tag{6.49}$$

At the equilibrium potential the following equality holds:

$$i^{\circ}_{\text{a, cond}} = i^{\circ}_{\text{c, cond}} = i^{\circ}_{\text{cond}} \tag{6.50}$$

With (6.43) and (6.48), we thus obtain:

$$i = i^{\circ}_{\text{cond}} - i^{\circ}_{\text{cond}} \exp\left(-\frac{F\eta}{RT}\right) \tag{6.51}$$

This equation provides a relation between the current density and the potential for an electron transfer reaction at an *n-type* semiconductor electrode.

For a *p-type semiconductor*, involving the transfer of electron holes between ferric ions and the valence band (with subscript v), an analogous derivation yields:

$$i_{\text{a}, v} = i^{\circ}_{\text{a}, v} \exp\left(\frac{F\eta}{RT}\right) \tag{6.52}$$

$$i_{\text{c}, v} = -i^{\circ}_{\text{c}, v} \tag{6.53}$$

$$i = i^{\circ}_{v} \exp\left(\frac{F\eta}{RT}\right) - i^{\circ}_{v} \tag{6.54}$$

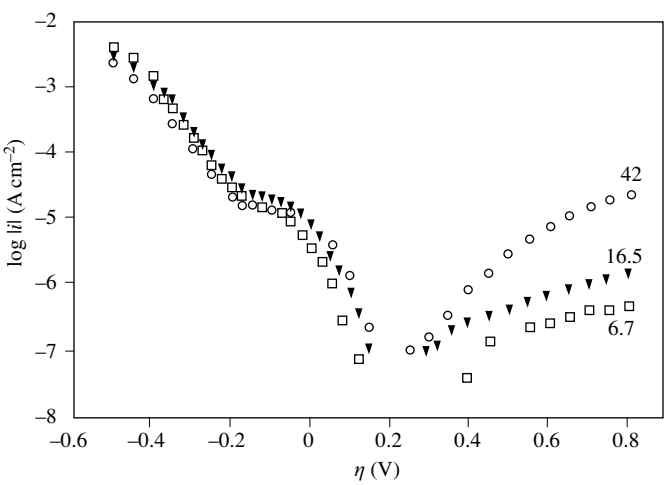


Figure 6.33 Measured polarization curves for the reaction $Fe(CN)_6^{3-} + e = Fe(CN)_6^{4-}$ on TiO_2 films of different donor concentration, $c_{\text{n,b}} = 42 \times 10^{18}$, 16.5×10^{18}, 6.7×10^{18} cm^{-3} [31].

© 2007, first edition, EPFL Press

Here, $i_{a,v}$ and $i_{c,v}$ are, respectively, the anodic and cathodic partial current densities. The variables $i^{\circ}_{c,v} = i^{\circ}_{a,v} = i^{\circ}_{v}$ represent kinetic parameters that are independent of the potential. Figure 6.31 shows dimensionless polarization curves calculated with the equations (6.51) and (6.54), respectively. The behavior is comparable to that of a diode that conducts current in a single direction.

Figure 6.32 represents expression (6.51) for an *n*-type semiconductor electrode in dimensionless form on a logarithmic scale. The calculated curve can be compared to the experimental data shown in Figure 6.33. They were obtained with the $Fe(CN)_6^{3-}/Fe(CN)_6^{4-}$ system on titanium oxide electrodes (an *n*-type semiconductor) having different concentrations of charge carriers. The measured anodic and cathodic current density-potential curves resemble those theoretically predicted. In general, charge transfer reactions at semiconducting oxide films do not exactly follow the simple theory outlined above, because other, more complex phenomena can also play a role.

Gerischer diagrams

Gerischer diagrams help to visualize the energy state of the semiconductor-electrolyte interface. These diagrams schematically present the band bending in the solid, and the density of states functions D_{ox} and D_{red} in the electrolyte. Figures 6.34 and 6.35 show Gerischer diagrams for an *n*-type semiconductor. At equilibrium (Figure 6.34), the Fermi level of the semiconductor is equal to the electrochemical potential of the electrons in the electrolyte. It therefore varies according to the

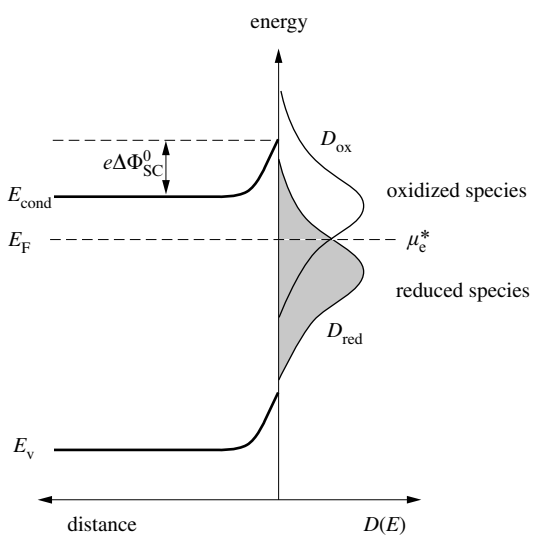


Figure 6.34 Semiconductor energy levels and density of state distribution function of dissolved species in contact with *n*-type semiconductor electrode at equilibrium.

© 2007, first edition, EPFL Press

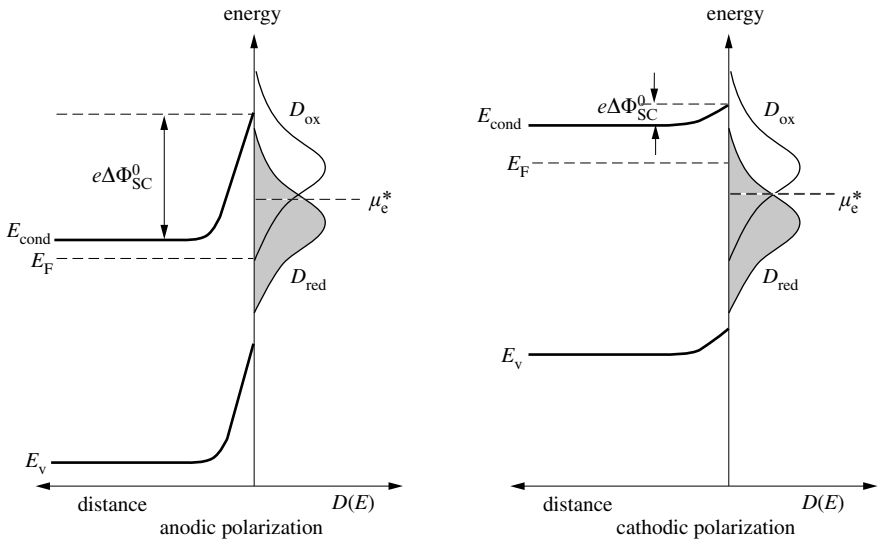


Figure 6.35 Effect of anodic and cathodic polarization on energy levels in *n*-type semiconductor electrode [32].

properties of the redox couple. The anodic and cathodic partial current densities at equilibrium are equivalent. Application of an anodic potential to the electrode (Figure 6.35(a)) moves the Fermi level and increases the degree of band bending. In this case, the electron concentration at the surface, n_s, decreases according to (6.43), and so does the cathodic partial current density. The result is therefore an overall anodic current. A cathodic overpotential (Figure 6.35(b)) causes the opposite effect.

According to Figure 6.34, electron exchange between the semiconductor and the ions in solution is facilitated when the distance between the conduction band and the Fermi level of the redox system is small. This is because the charge transfer rate depends on D_{ox} and D_{red}. Generally, the rate of electron exchange will be more or less rapid, depending on the reversible potential of the dissolved species and on the band structure of the oxide.

6.3 PASSIVITY BREAKDOWN AND TRANSPASSIVE BEHAVIOR

6.3.1 Anodic dissolution in the transpassive potential region

At the transpassivation potential E_b (Figure 6.3), the properties of the passive film change and one observes a renewed increase in the rate of dissolution. This behavior is referred to as ***anodic depassivation***. It may be the result of film oxidation at high anodic potentials or of film breakdown favored by the presence of certain anions. Generally speaking, in the transpassive potential region, one observes three types of metal dissolution behavior:

© 2007, first edition, EPFL Press

- uniform transpassive dissolution at potentials below oxygen formation;
- localized dissolution by pitting;
- high rate transpassive dissolution above the potential of oxygen formation.

Uniform transpassive dissolution by film oxidation

Uniform transpassive dissolution below the potential for oxygen formation occurs when the protecting passive film is oxidized to species having a higher solubility. This mechanism is typically observed with metals of column VI of the periodic table, chromium, molybdenum and tungsten, which can form ***soluble oxoanions*** of valence +6. We shall discuss this mechanism for chromium, which is the most important and best known example of this type of behavior. Upon passivation, chromium forms a passive film of Cr_2O_3, which is only sparingly soluble and provides an effective barrier that protects the metal against corrosion. At the transpassivation potential, the trivalent chromium oxide is oxidized to hexavalent oxide, CrO_3, which has a high solubility in aqueous media, where depending on pH it forms chromate ions CrO_4^{2-} (basic conditions) or bichromate ions $Cr_2O_7^{2-}$ (acidic conditions). The transpassive dissolution of chromium therefore corresponds to the following stoichiometric equations:

$$CrO_4^{2-} + 4\,H_2O + 6\,e \;=\; Cr + 8\,OH^- \qquad E^\circ = 0.37\ \text{V} \tag{6.55}$$

$$Cr_2O_7^{2-} + 14\,H^+ + 12\,e \;=\; Cr + 7\,H_2O \qquad E^\circ = 0.29\ \text{V} \tag{6.56}$$

The standard potentials of these two electrode reactions are significantly lower than that of the oxygen electrode. The transpassive dissolution of chromium therefore can occur at potentials well below those needed for oxygen evolution. As an illustration, Figure 6.36 shows anodic polarization curves for the transpassive

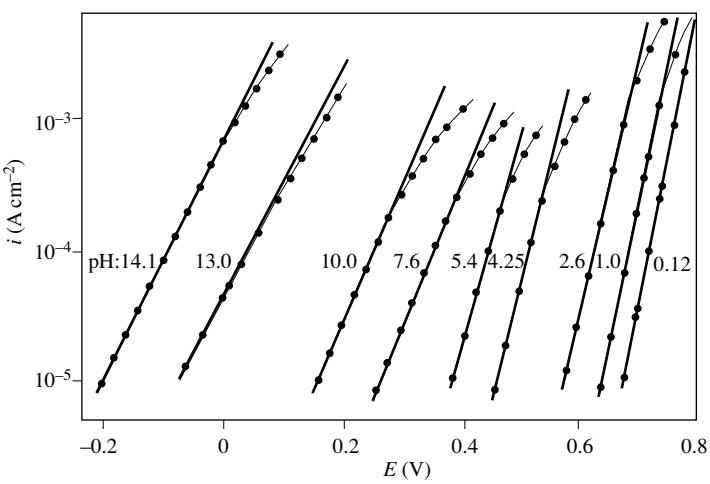


Figure 6.36 Polarization curves for chrome in the transpassive region at different pH [33].

© 2007, first edition, EPFL Press

dissolution of chromium at different pH [33]. With decreasing pH the Tafel lines shift to higher potentials, because the reversible potential also shifts with pH. The transpassive dissolution of chromium, in general, is uniform in a similar way as active dissolution.

The oxidation of trivalent chromium oxide to hexavalent oxide is also observed in passive films on stainless steel. An interesting application of this effect is ***electropolishing.*** The object made of stainless steel to be polished is connected as the anode in an electrochemical cell that contains a suitable electrolyte, normally a mixture of concentrated sulfuric and phosphoric acids. Transpassive dissolution then leads to electropolishing, provided the dissolution reaction is under mass transport control. Protrusions on the surface are favored by mass transport and therefore they dissolve more rapidly, leading to leveling of the surface [34].

Figure 6.37 shows anodic polarization curves for an iron-chromium alloy (Fe-13Cr) in concentrated sulfuric-phosphoric acid measured with a rotating disk electrode at different rotation rates. Well defined limiting current plateaus are observed in the transpassive potential region, their height increasing with rotation rate. A plot of $1/i_l = f\ (1/\omega^{1/2})$ shown in the insert produces a straight line passing through the origin. The anodic limiting current density therefore is entirely controlled by mass transport. The current increase at the end of the limiting current plateau is due to oxygen evolution, which in this potential region can take place in addition to metal dissolution.

Localized dissolution by pitting

In the presence of certain anions, in particular chloride, metals lose their passivity above a critical potential, called the ***pitting potential,*** and metal dissolution takes place

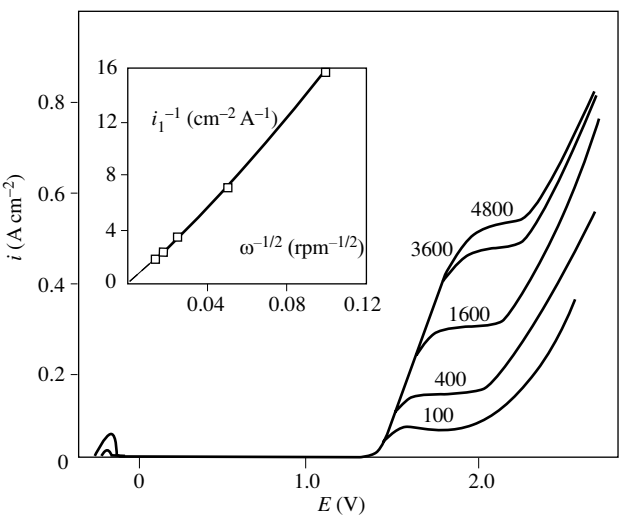


Figure 6.37 Anodic polarization curves for an iron-chromium alloy (Fe-13Cr) in an electrolyte consisting of 65% H_3PO_4, 20% H_2SO_4 at 70 °C, measured on a rotating disk electrode at different rotation rates: 100, 400, 1600, 3600, 4800 rpm [35].

© 2007, first edition, EPFL Press

from local sites where the passive film broke down. The localized dissolution leads to formation of deep pits on an otherwise passive surface. The described phenomenon is responsible for ***pitting corrosion*** that we shall discuss in more detail in Chapter 7. The size of the pits formed varies; their diameter can reach several tens of micrometer. Their depth can be of similar magnitude, or even greater. Figure 6.38 shows pits formed on an iron-chromium alloy (Fe-24Cr) by anodic polarization above the pitting potential in 1M HCl solution.

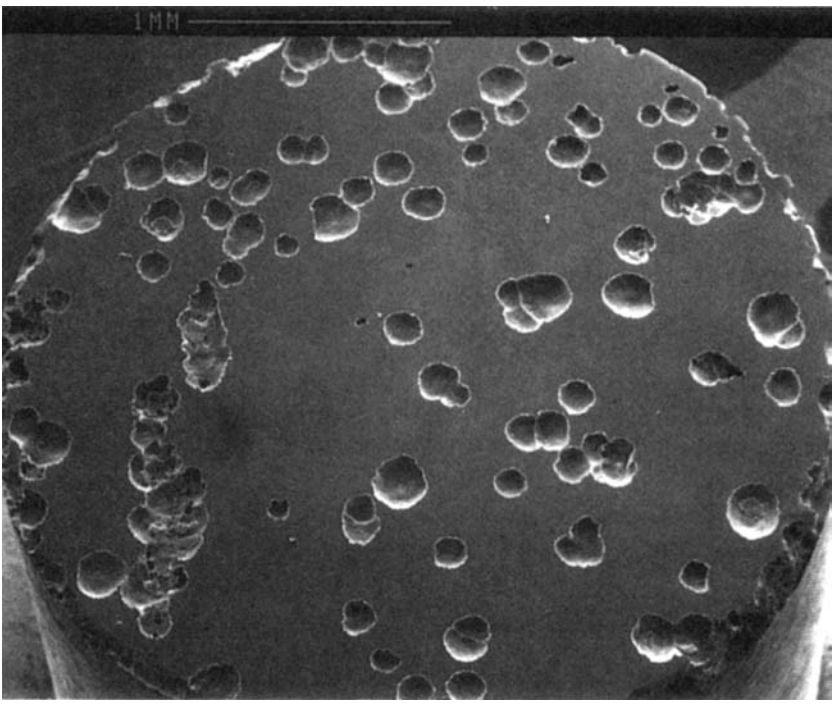


Figure 6.38 Surface pitting of an Fe-24Cr alloy resulting from anodic polarization in 1M HCl above the pitting potential.

The depassivation mechanism responsible for pitting involves specific effects of certain electrolyte anions, called ***aggressive anions***, as will be discussed in the next section.

High-rate transpassive dissolution

In absence of aggressive anions (for example in a sodium nitrate solution) the transpassivation potential of metals such as iron and nickel which form electron conducting oxides is higher than the reversible potential of the oxygen electrode: $E_b > E_{rev,O2}$. As a result, when such a metal is anodically polarized to sufficiently high potentials, anodic oxygen evolution is observed. Eventually, as the film breakdown potential is reached, oxygen evolution gives way to anodic dissolution at high rate. The passive film breakdown is favored on one hand by a high anion concentration and

© 2007, first edition, EPFL Press

on the other hand by the high local acidity at the anode surface which is the result of oxygen formation from water, which liberates protons.

Figure 6.39 illustrates the described behavior for the transpassive dissolution of nickel in an alkaline solution of sodium nitrate [36]. The anode potential, the current efficiency for metal dissolution and the apparent thickness of the passive film measured by coulometry are presented as a function of applied current density. For low current densities (relatively low transpassive potential region), oxygen formation is the dominating reaction and therefore the current efficiency for metal dissolution is almost zero. With increasing current density (increasing potential) the rate of metal dissolution becomes more important and eventually accounts for most of the current. This is evidenced by the value of the current efficiency for metal dissolution, which reaches almost 100 %. The thickness of the passive film goes through a maximum and sharply falls when metal dissolution sets in. In this system the anion concentration plays a decisive role for anodic depassivation, the transpassive metal dissolution being favored by a high nitrate concentration at the electrode-solution interface [37].

High rate transpassive dissolution at potentials exceeding that of oxygen formation is never observed when a metal corrodes in a natural environment, because under these conditions oxygen is reduced and the mixed potential of the corroding metal lies always below that of the oxygen electrode. On the other hand, high-rate transpassive dissolution is used in the ***electrochemical machining*** of metals: a pre-shaped tool is connected as the cathode and a workpiece as the anode, the distance between the two

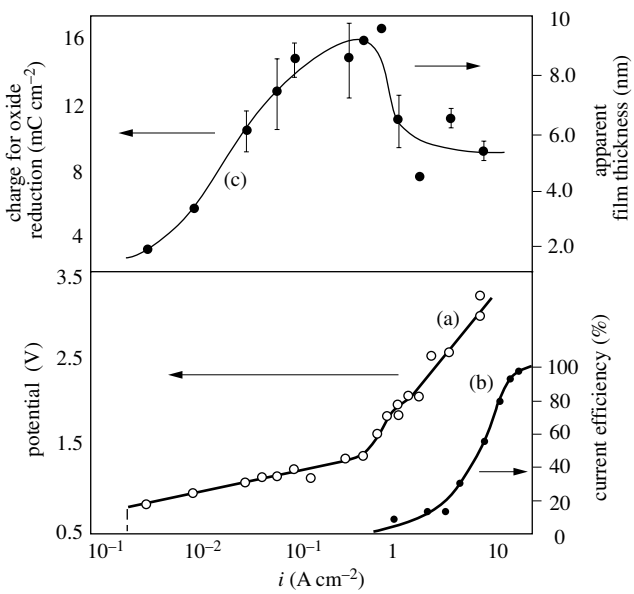


Figure 6.39 Galvanostatic polarization of Ni in the transpassive potential region in sodium nitrate solution: (a) anodic polarization curve in the transpassive potential region; (b) the current efficiency for metal dissolution; and (c) the apparent thickness of the film as measured by coulometry [36].

© 2007, first edition, EPFL Press

electrodes being just some fraction of millimeter. During dissolution, the cathode is advanced so that the inter-electrode distance remains constant. The anodic dissolution thus reproduces the shape of the cathode onto the anode. Rapid electrolyte circulation eliminates the reaction products, and the heat produced by the Joule effect at current densities on the order of 20–200 A cm^{-2}. Transpassive dissolution also serves for electrochemical micromachining, where the metal is anodically dissolved through a patterned photoresist [38].

6.3.2 Passive film breakdown and pit initiation

The mechanism of passivity breakdown leading to dissolution by pitting has been widely studied, because of its importance for the development of alloys with improved resistance to pitting corrosion. Unfortunately, despite a considerable amount of research done, we still do not fully understand the mechanisms involved in film breakdown and pit initiation. Indeed, there are a number of potential causes for the local loss of passivity, when a metal is polarized to a high potential in presence of aggressive anions. On industrial quality alloys, selective dissolution of inclusions may take place above a certain potential that can lead to the formation of pits (Chapter 7). On metals subjected to elastic or plastic deformation, mechanical rupturing of the passive film may be the source for local attack. But local depassivation and pitting occur also on pure metals not subjected to mechanical forces. Figure 6.40 presents three possible mechanisms that have been proposed to explain this behavior:

- anion penetration into the passive film;
- depassivation at defects;
- film dissolution promoted by anion adsorption.

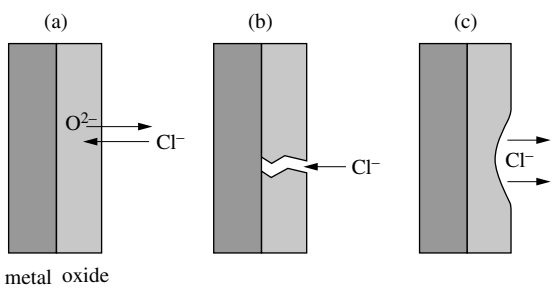


Figure 6.40 Proposed anodic depassivation mechanisms in presence of chloride: (a) anion penetration into the film by exchange with O^{2-}; (b) electrolyte penetration at the metal surface through film defects; (c) film dissolution caused by local anion adsorption.

Depassivation due to anion penetration

According to this theory, aggressive anions such as chloride penetrate into the passive film, under the influence of the electric field, where they exchange positions in the crystal lattice with O^{2-} ions (Figure 6.40(a)). This process is believed to modify the film properties, such as increasing its conductivity and facilitating its dissolution.

© 2007, first edition, EPFL Press

To explain the non-uniformity of the film breakdown it has been proposed that exchange of divalent oxygen anions with monovalent chloride creates vacancies in the film lattice, which migrate to the film-metal interface and coalesce there [39]. The presence of chloride ions in passive films has indeed been reported by some authors using surface analysis. However, several other studies indicate that although anions may be incorporated during film growth, a simple exposure of a stable passive film to chloride does not lead to ion exchange and penetration of chloride into the film.

When a constant potential above the pitting potential is applied to a passive metal, pitting does not take place immediately, but only after a certain time. This ***induction time*** for pit initiation becomes smaller at higher applied potentials and with increasing concentration of aggressive anions. A possible migration of electrolyte anions from the exterior surface of the film towards the film-metal interface is qualitatively consistent with this behavior; it takes some time for the migrating anions to reach the metal surface. At this time, no comparison between theoretical and experimental time constants for chloride penetration involving migration through the oxide film lattice is available to test this mechanism. Indeed, one may expect that the rate of penetration of ions through a passive film would depend mostly on the presence of structural defects, and therefore be rather irreproducible.

Depassivation at defects

A number of authors have attributed anodic depassivation in presence of aggressive anions to the presence of pre-existing defects in the passive film, such as pores or cracks formed during its growth (Figure 6.40(b)) or imperfections related to the substrate structure. At such defects the electrolyte may enter into direct contact with the metal surface, leading locally to rapid dissolution and formation of a pit. Experience shows that high purity metals with a well defined surface structure pit less readily. Passive films formed on such substrates contain fewer imperfections, and are therefore less prone to breakdown and pit initiation.

Cracks are another type of defect that could be the location of pit initiation. Cracks can form in a brittle oxide film due to plastic deformation of the underlying metal. For example, the brittle oxide film formed by anodizing aluminum crack quite easily. Some researchers have proposed that the strong electric field present in passive films produces an electrostrictive force sufficiently strong to break up a passive film. Experimental evidence to sustain this hypothesis is lacking, however.

Whatever the origin of defects in a passive film, a local loss of passivity can only occur when the exposed metal surface does not immediately repassivate. Indeed, it has been observed that already well below the critical pitting potential depassivation and repassivation events may occur. These can be seen particularly well when working with electrodes of small surface area (microelectrodes), because they contain relatively few defects that lead to breakdown events. Individual events therefore can be studied more easily. The results of Figure 6.41 illustrate the described behavior. It presents potentiostatic transients observed in the passive potential region on an iron-chromium alloy in NaCl using a microelectrode [40]. Each individual current peak represents a

© 2007, first edition, EPFL Press

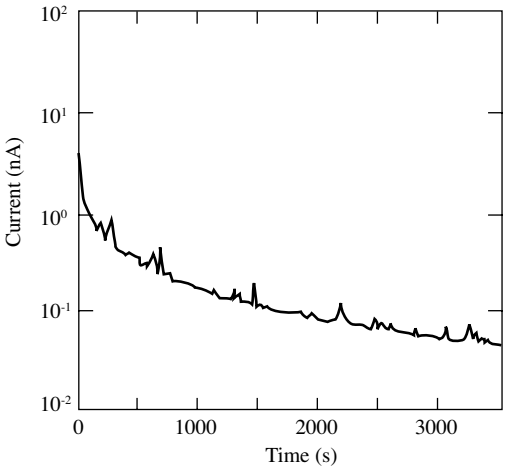


Figure 6.41 Variation of current with time measured on a microelectrode of 0.2 mm diameter during anodic polarization of an Fe-25Cr alloy in the passive potential region in 1 M NaCl. The current spikes reflect local depassivation/repassivation events [40].

depassivation and repassivation event. Their number and intensity varies with alloy composition and potential. Sustained dissolution by pitting may occur once a critical potential is exceeded and the ***rate of repassivation*** becomes insufficient. The role of aggressive anions is to favor metal dissolution at the expense of oxide formation at defects sites where the electrolyte enters into in contact with the metal surface.

Film dissolution promoted by anion adsorption

The scheme shown in Figure 6.40(c) postulates that anodic depassivation is caused by an accelerated dissolution of the passive film as a result of adsorption of aggressive anions at the outer film surface, replacing hydroxyl groups. Evidence for such a behavior in case of iron has been provided by an experiment using a ***rotating ring-disk electrode*** [41]. In this arrangement a rotating disk electrode is surrounded by a concentric ring electrode that is electrically isolated from the disk (Figure 6.42). Rotation causes a convective motion that carries the reaction products from the disk towards the ring. By applying an appropriate potential to the ring, it is possible to oxidize or reduce a reaction product formed at the disk. In many cases, a quantitative analysis of reaction products can be made, because the theoretical collection efficiency at the ring depends only on geometry [42].

The application of the rotating ring-disk technique to anodic depassivation of iron is based on the following principle. When the iron disk is maintained at an anodic potential in the passive region, film dissolution produces small amounts of Fe^{3+} ions. By setting the potential at the ring at a value below the reversible potential of the Fe^{3+}/Fe^{2+} couple, the Fe^{3+} ions that arrive at the ring are reduced to Fe^{2+}. Alternatively,

© 2007, first edition, EPFL Press

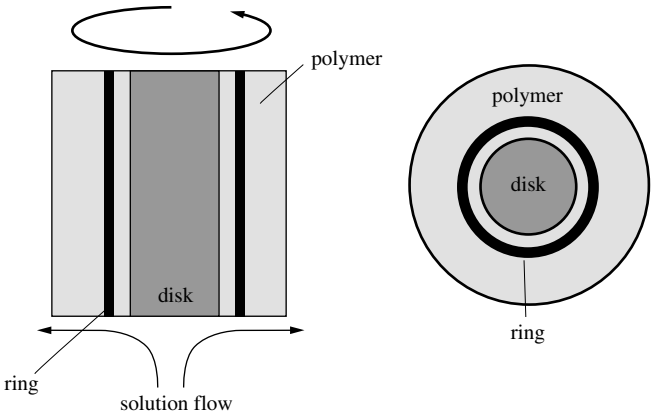


Figure 6.42 Rotating ring-disk electrode (schematic).

one could measure the rate of Fe^{2+} formation at the disk by applying a potential to the ring that oxidizes Fe^{2+} ions into Fe^{3+}.

In the experiment of Figure 6.43, the iron disk was anodically polarized by applying a potential located in the passive region. When the current density reached a

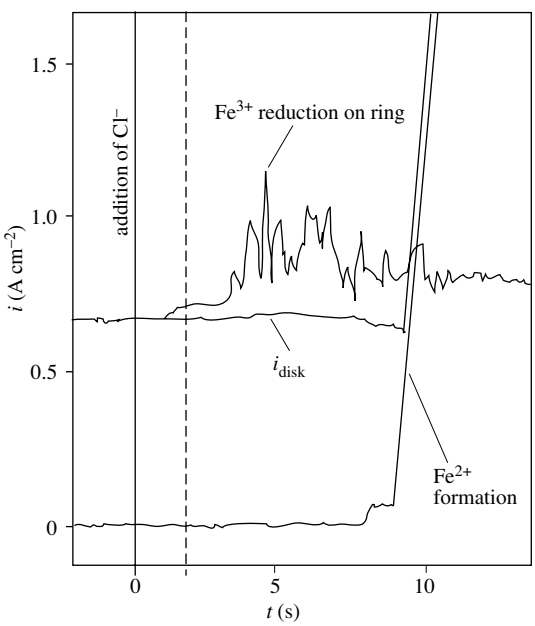


Figure 6.43 Rotating ring-disk experiment of passive film breakdown on Fe after injection of chloride ions. Parameters observed to vary as a function of time: the current density of the disk; the reduction current density of Fe^{3+} on the ring; and the current density of formation of Fe^{2+} on the disk. Electrolyte is borate solution of pH 7.3. Concentration of NaCl added at time t = 0 is 0.01mol l^{-1}. Dashed line shows the time to reach a uniform NaCl concentration [41].

© 2007, first edition, EPFL Press

steady-state, chloride ions were added to the solution while maintaining the potential. With the addition of chloride, one observes an increase in the current for Fe^{3+} reduction at the ring. This indicates that the dissolution rate of the passive film has increased. After some time, corresponding to pit initiation, the total current at the disk rises sharply, and so does the current for Fe^{2+} formation, calculated from the difference of total current and Fe^{3+} current. The fact that most of the current at the disk serves for Fe^{2+} formation is a clear sign of dissolution by pitting. Indeed, it is well known that iron dissolution from pits produces Fe^{2+}, in a similar way as active dissolution, while passive dissolution produces Fe^{3+} ions.

Pitting means that dissolution is located on a few spots on an otherwise passive surface. To explain why in the foregoing experiments dissolution is localized rather than uniform, one has to postulate that anion adsorption is not uniform, but occurs at only certain sites. At these sites the adsorbed anions increase locally the rate of dissolution, causing local thinning of the film. This in turn increases the local electric field and thus further accelerates the rate of film dissolution until a pit is formed and the metal enters in direct contact with the electrolyte.

The thickness of passive films at steady state depends on the applied potential and their rate of dissolution (equation 6.37). If chloride ions accelerate the rate of dissolution, a higher chloride concentration in the electrolyte should lead to thinner passive films. This behavior has been confirmed by XPS measurements, as shown in Figure 6.44. This figure presents the effect of several halogen-anions on the thickness of the passive films formed on iron in a phthalate buffer solution (pH 5). In this case, Cl^- ions have a greater effect than either Br^- or I^- ions.

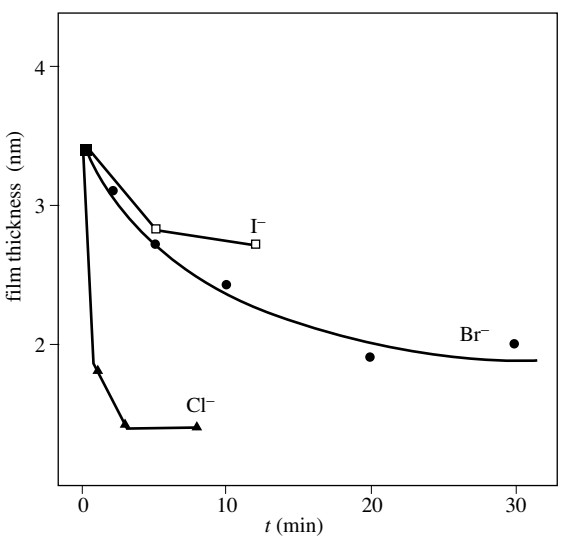


Figure 6.44 Thickness of iron passive film measured by XPS after different exposure times to solutions containing 0.1 mol/l of Cl^-, Br^-, or I^- anions. Applied potential is $E = 0.5$ V. The films were initially formed in a phthalate solution of pH 5.0 at 1.2 V [43].

© 2007, first edition, EPFL Press

Overall assessment of proposed film breakdown mechanisms

The structure of surfaces of metals and alloys used in engineering is never perfect (see Chapter 3). These surfaces reproduce structural features of the bulk metal such as phase boundaries, grain boundaries, emerging dislocations, non metallic inclusions of different composition and size, they exhibit roughness at different scales, and they are often plastically deformed due to machining or polishing. Thin passive oxide films are able to cover up many of these imperfections, but it is unlikely that locally their structure is not affected by them. Passive films on real metal surfaces therefore contain a variety of structural defects that can act as preferred sites for anodic depassivation and pit initiation.

We have seen that pit initiation by local loss of passivity requires the presence of certain anions. Any explanation of pit initiation therefore must consider how the chemical properties of aggressive anions can affect their interaction with passive metals. The adsorption behavior of different anions on the passive oxide film is an important aspect, and how it affects film dissolution on an atomic scale. Another critical question concerns the effect of anions on the relative rate of dissolution and repassivation. Once the electrolyte enters into contact with the metal surface at a defect site, two types of anodic reactions are in competition; formation of oxide species and of hydrated or complexed metal ions. Depending on which is faster, repassivation or dissolution dominates in a pit nucleus. Both anodic processes involve intermediate reaction steps and adsorbed species. The specificity of aggressive anions thus is most likely the result of their adsorption behavior on the metal surface and their participation in the anodic reaction mechanisms by formation of reaction intermediates. Interestingly, at sufficiently high potentials all types of anions apparently can lead to anodic depassivation of metals. However, for certain anions such as nitrate ions, the critical potential for film breakdown lies well above the potential for oxygen formation and therefore, under normal corrosion conditions, passive films are stable against pitting in these environments.

In conclusion, the present discussion of proposed film breakdown and pit initiation mechanisms suggests that several phenomena are responsible for the loss of passivity and the onset of pitting when a metal is polarized to a high potential in presence of aggressive anions. Structural defects in the passive film reflecting those of the metal, anion adsorption on the film and the metal surfaces and the effect of anions on the kinetics of the electrochemical reactions governing oxide formation and metal dissolution are most critical. Practical consequences of these phenomena for pitting corrosion will be discussed in Section 7.3.

General bibliography

R. B. Frankenthal and J. Kruger, eds., *Passivity of Metals*, The Electrochemical Society, Inc., Princeton N.J. 1978, 1077 pp.

B. Ives J. L. Lou, J. R. Roa, eds., *Passivity of Metals and Semiconductors*, Proc. Vol. 99-42, The Electrochemical Society, Inc., Pennington N.J. (2001) 974 pp.

© 2007, first edition, EPFL Press

P. Marcus, Editor, *Corrosion Mechanisms in Theory and Practice*, 2nd edition, Marcel Dekker, New York, 2002, 742 pp.
P. Marcus and F. Mansfeld, Editors, *Analytical Methods in Corrosion Science and Engineering*, CRC Press, Boca Raton, 2005, 776 pp.

References

[1] P. Marcus, I. Olefjord, *Surf. Interf. Analysis* 11, 569 (1988).
[2] U. L. Heubner, E. Altpeter, M. B. Rochel, E. Wallis, *Corrosion NACE*, 45, 254 (1989).
[3] A. J. Bard, R. Parsons, J. Jordan, *Standard Potentials in Aqueous Solution*, Marcel Dekker, New York (1985) p.1–833.
[4] Y. M. Kolotyrkin, *Stress Corrosion Cracking and Hydrogen Embrittlement of Iron Base Alloys,* NACE 5, Natl. Ass. Corrosion Engineers, Houston, Texas (1977) p. 946.
[5] K. Vetter, *Electrochemical Kinetics*, Academic Press, New York (1967) p. 751.
[6] G. Okamoto, *Corr. Sci.* 13, 471 (1973).
[7] H. H. Uhlig, R. W. Revie*, Corrosion and Corrosion Control,* third edition, John Wiley, New York, (1985) p. 78-79.
[8] P. Süry, *Corr. Sci.* 16, 879 (1976).
[9] E. A. Lislovs, A.P. Bond, *J. Electrochem. Soc.*, 122 719 (1975).
[10] P. Russel, J. Newman, *J. Electrochem. Soc.* 134, 1051 (1987).
[11] N. Sato, *Comprehensive Treatise of Electrochemsitry*, vol. 4, Plenum Press, New York, (1981) p. 197-198.
[12] P. Marcus, *Electrochim. Acta* 43, 109 (1998).
[13] C. L. Mc Bee, J. Kruger, *Electrochim. Acta* 17, 1337 (1972).
[14] M. P. Ryan, R. C. Newman, G. E. Thompson, *J. Electrochem. Soc.* 142, L177 (1995).
[15] D. Zuili, V. Maurice, P. Marcus, *J. Phys. Chem. B* 103, 7896 (1999).
[16] H. Kaesche, *Die Korrosion der Metalle*, third edition, Springer Verlag, Heidelberg (1990) p. 206.
[17] K. Asami, K. Hashimoto, S. Shimodaira, *Corr. Sci.* 18, 151 (1978).
[18] D. Landolt, *Advances in Localized Corrosion*, NACE-9, H. Isaacs, U. Bertocci, J. Kruger, S. Smialovska, editors, Natl. Assoc. Corrosion Engineers., Houston, Texas (1990) p. 25.
[19] H. Fischmeister, U. Roll, *Fresenius Zeitschrift f. anal. Chemie* 319, 639 (1984).
[20] C. Calinski, H.H. Strehblow, *J. Electrochem. Soc.* 136, 1328 (1989).
[21] P. Schmutz, D. Landolt, *Corr. Sci.* 41, 2143 (1999).
[22] D. A. Vermilyea, *Act. Met.* 1, 282 (1953).
[23] N. Cabrera, N.F. Mott, *Rep. Prog. Phys.*, 12, 163 (1948/49).
[24] S. Haupt, C. Calinski, U. Collisi, H.W. Hoppe, H.D. Speckmann, H.H. Strehblow, *Surf. Interf. Analysis* 9, 357 (1986).
[25] K. J. Vetter, F. Gorn, *Electrochim. Acta* 18, 321 (1973).

© 2007, first edition, EPFL Press

[26] R. Kirchheim, *Electrochim. Acta* 32, 1619 (1987).
[27] L. Zhang, D.D. Macdonald, E. Sikora, J. Sikora, *J. Electrochem. Soc.* 145, 898 (1998).
[28] C.-A. O. Olsson, D. Hamm, D. Landolt, *J. Electrochem. Soc.* 147, 4093 (2000).
[29] C.-A. O. Olsson, M.-G. Vergé, D. Landolt, *J. Electrochem. Soc.* 151, B652 (2004).
[30] G. Okamoto, T. Shibata, R. B. Frankenthal, J. Kruger, Editors, Passivity of Metals, The Electrochemical Society, Inc., Princeton N.J. (1978) pp. 646, 656.
[31] W. Schmickler, J. W. Schultze, *Modern Aspects of Electrochemistry,* No. 17, J.O'M Bockris, B. E. Conway, R. White, eds., Plenum Press, New York (1986) p. 357.
[32] A. J. Bard, R. Memming, B. Miller, *Pure and Appl. Chem.* 63, 569 (1991).
[33] K. E. Heusler, *Passivity of Metals*, R. B. Frankenthal, J. Kruger, Editors, The Electrochemical Society, Inc., Princeton N.J. (1978) p.77.
[34] D. Landolt, *Electrochim. Acta* 32, 1 (1987).
[35] L. Ponto, M. Datta, D. Landolt, *Surf. Coatings Technol.* 30, 265 (1987).
[36] M. Datta, D. Landolt, *Electrochim. Acta* 24, 843 (1979).
[37] D. Landolt, *Werkstoffe und Korrosion* 45, 188 (1973).
[38] M. Datta M., L. T. Romankiw, *J. Electrochem. Soc.* 136, 285C (1989).
[39] C. Y. Chao, L. F. Lin, D. D. Macdonald, *J. Electrochem. Soc.* 128, 1191 (1981).
[40] Y. Kobayashi, S. Virtanen, H. Bohni, *J. Electrochem. Soc.* 147, 155 (2000).
[41] K. E. Heusler, L. Fischer, *Werkstoffe und Korrosion* 27, 551 (1976).
[42] W. J. Albery, *Trans. Farad. Soc.* 62, 1915(1966).
[43] W. Khalil, S. Haupt, H. H. Strehblow, *Werkstoffe und Korrosion* 36, 16 (1985).

© 2007, first edition, EPFL Press

CHAPTER 7

LOCALIZED CORROSION PHENOMENA

7.1 CORROSION CELLS

7.1.1 Origin and properties of corrosion cells

Whenever two parts of a metallic structure that is in contact with an electrolyte are at a different potential, they form an electrochemical cell, called ***corrosion cell***. Potential differences can result from local differences in the material (different metals, phases, grain boundaries, inclusions, segregations) or from differences in the local environment (exposure to air, pH, convection, temperature). Table 7.1 gives a few examples.

Table 7.1 Origin of corrosion cells.

Origin of potential difference	Designation
Contact between two different metals	Galvanic corrosion
Difference in oxygen accessibility	Differential aeration cell, Crevice corrosion
Grain boundary precipitation	Intergranular corrosion
Local depassivation by aggressive anions	Pitting corrosion
Difference in temperature	Corrosion due to thermal cells

Whatever their origin, corrosion cells share a number of common features which we shall discuss by using a very simple model. We represent the corrosion cell by an electrochemical cell (Figure 7.2) consisting of two flat electrodes of same surface area in contact with an electrolyte, and we assume that the current density is uniform throughout.

At open circuit (Figure 7.2(a)), the corrosion potential of the electrode (I) is equal to $E_{cor,I}$ and that for electrode II is equal to $E_{cor,II}$. Let us assume that $E_{cor,I} < E_{cor,II}$. As soon as we close the circuit between the electrodes and thus bring them into electrical contact (Figure 7.2(b)), a current I, called the external current, flows from the positive pole (II) to the negative pole (I). The electrons circulate in the opposite direction. The two electrodes undergo a polarization ζ, positive at the anode (I) and negative at the cathode (II), thus narrowing the potential difference between them:

$$E_I = E_{cor,I} + \zeta \text{ and } E_{II} = E_{cor,II} + \zeta \tag{7.1}$$

© 2007, first edition, EPFL Press

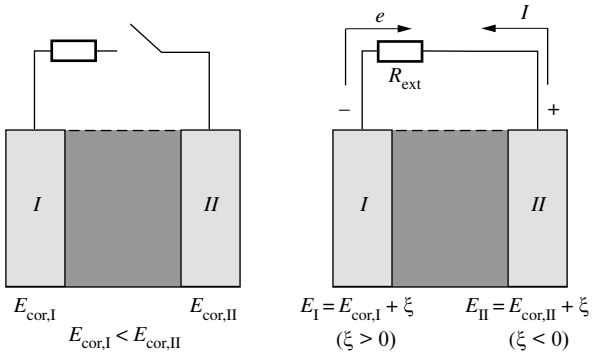


Figure 7.2 Corrosion cell due to contact of two different metals: (a) open circuit and (b) closed circuit configurations. R_{ext} corresponds to the external resistance.

Figure 7.3 represents the behavior of such a cell in an Evans diagram (plotting the logarithm of the absolute value of the current as a function of the potential). We can write for the net current at the electrode (I):

$$I_I = I_{a,I} + I_{c,I} \tag{7.2}$$

Similarly for electrode (II):

$$I_{II} = I_{a,II} + I_{c,II} \tag{7.3}$$

The net current at electrode (II) takes on a negative sign because it corresponds to a cathodic reaction. The external current flow is thus equal to $I = I_I = -I_{II}$. The principle of charge conservation requires that the sum of all partial currents is zero:

$$I_{a,II} + I_{c,II} + I_{a,I} + I_{c,I} = 0 \tag{7.4}$$

When the two electrodes are brought into contact, the anodic partial current, proportional to the corrosion rate, increases at electrode (I). Simultaneously, it

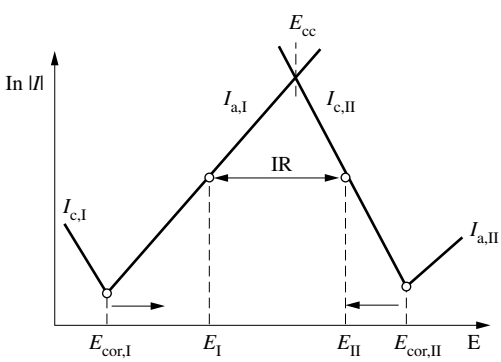


Figure 7.3 Evans diagram of a corrosion cell.

© 2007, first edition, EPFL Press

decreases at electrode (II). Conversely, the cathodic partial current decreases at electrode (I) and increases at electrode (II). In corrosion cells, the anodic current at electrode (II), as well as the cathodic current at electrode (I), are often negligibly small. We can then simplify (7.4), which gives:

$$I = I_{a,I} = -I_{c,II} \tag{7.5}$$

Formally, a corrosion cell can thus be treated as a mixed electrode just as in uniform corrosion, but there is one important difference: the anodic and cathodic surfaces are not necessarily equal. As a consequence, if according to (7.5) the *intensity* of anodic and cathodic partial currents are always equivalent, the respective *current densities* may differ depending on the surface area ratio between cathode and anode.

According to Faraday's law the anodic partial current density is proportional to the rate of corrosion. Thus, if the currents $I_{c,I}$ and $I_{a,II}$ are negligible, we get from equation (7.5) for the corrosion rate v_I of the anode:

$$v_I = \frac{i_{a,I}}{n_I F} = -\frac{i_{c,II}}{n_I F}\frac{A_{II}}{A_I} \tag{7.6}$$

In this equation, we have set $i_{a,I} = I_{a,I}/A_I$ and $i_{c,II} = I_{c,II}/A_{II}$. It follows that in a corrosion cell, the rate of corrosion of the anode is the higher the smaller its surface area is with respect to that of the cathode (***surface area effect***).

We can draw an electrical circuit to represent a corrosion cell, as shown in Figure 7.4. The difference in corrosion potential between the two electrodes represents a voltage source that has the value of $\Delta E_{cor} = E_{cor,II} - E_{cor,I}$. The polarization resistance at each of the two electrode-electrolyte interfaces is represented by a non-ohmic resistive element: $R_{p,I} = (dE/dI)_I$ and $R_{p,II} = (dE/dI)_{II}$. The value of these circuit elements varies with the current (Sect. 4.2). The ***internal resistance*** R_{int} represents the ohmic resistance of the electrolyte between the two electrodes, and the ***external resistance*** R_{ext} represents the resistance arising from the electronic conductors of the circuit.

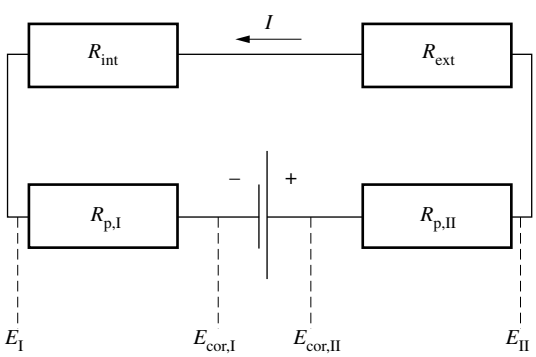


Figure 7.4 Electrical equivalent circuit of corrosion cell.

© 2007, first edition, EPFL Press

According to Figure 7.4, the potential difference $\Delta E = E_{II} - E_I$ that is found in a corrosion cell is given by equation (7.7), where I represents the external current and R is the sum of the internal and external ohmic resistances, $R = R_{int} + R_{ext}$.

$$\Delta E = IR = I(R_{int} + R_{ext}) \tag{7.7}$$

If $(R_{int} + R_{ext}) = \infty$, then $I = 0$, and we find:

$$\Delta E = \Delta E_{cor} = E_{cor,II} - E_{cor,I} \tag{7.8}$$

This shows that a corrosion cell cannot exist unless two parts of a metallic structure acting as electrodes are connected by both an electrolyte and an electron conducting material. Otherwise, each electrode will corrode independently of the other.

According to (7.7), if the sum of the internal and external resistances is zero ($R_{int} + R_{ext} = 0$), then the potential difference also becomes zero and, as a result, the two electrodes are at the same potential ($E_I = E_{II}$). This means that a short circuit has been created. In Figure 7.3, E_{cc} indicates the potential at short circuit. In practice the two potentials E_I and E_{II} are never equal because the resistance of the electrolyte is not zero.

From Figure 7.4, it follows that the external current I of a corrosion cell is given by

$$I = \frac{\Delta E_{cor}}{R_{p,I} + R_{int} + R_{ext} + R_{p,II}} = \frac{E_{II} - E_I}{R_{int} + R_{ext}} \tag{7.9}$$

The current that flows between the anode and the cathode therefore depends on the value of all circuit elements shown. Unfortunately, in real situations, the value of these elements is usually not known precisely. The parameters $R_{p,I}$ and $R_{p,II}$ vary as a function of current density and depend on reaction kinetics. The value of R_{int} depends not only on electrolyte conductivity but also on the system geometry and as a consequence, the current densities in real corrosion cells may be highly non uniform. A quantitative analysis of corrosion cells therefore would require a full calculation of the potential distribution between the anodic and cathodic areas (Section 12.3).

To get a qualitative insight into the behavior of corrosion cells, we consider two limiting cases with respect to the reaction kinetics, assuming a uniform current density:

- ohmic control;
- control by oxidant transport.

Ohmic control

If the polarization resistances at the interfaces are small relative to the electrolyte resistance ($R_{p,I} << R_{int}$ and $R_{p,II} << R_{int}$) we can write $E_I \approx E_{cor,I}$ and $E_{II} \approx E_{cor,II}$. If in addition, the electrolyte resistance greatly exceeds the external resistance ($R_{int} >> R_{ext}$), equation (7.9) becomes:

© 2007, first edition, EPFL Press

$$I = \frac{\Delta E_{\text{cor}}}{R_{\text{int}}} \tag{7.10}$$

Usually in a corrosion cell, the external current I is approximately equal to the anodic current $I_{\text{a,I}}$. We find thus for the corrosion rate:

$$v_{\text{I}} = \frac{I_{\text{a,I}}}{n_{\text{I}} F A_{\text{I}}} = \frac{1}{n_{\text{I}} F A_{\text{I}}} \frac{\Delta E_{\text{cor}}}{R_{\text{int}}} \tag{7.11}$$

Control by oxidant transport

If the cathodic reaction (oxygen reduction) takes place at the limiting current, then the resistance at the electrode-solution interface of the cathode, $R_{\text{p,II}}$ (Figure 7.4), behaves as a non-linear electrical element that limits the current (current limiter). Indeed, at the limiting current plateau one has: $R_{\text{p,II}} = (\text{d}E/\text{d}I)_{\text{II}} = \infty$. The anode corrosion rate v_{i} depends, in this case, only on the value of limiting current density at the cathode, $i_{\text{l,II}}$ that is determined by prevailing mass transport conditions. Neglecting the currents $I_{\text{c,I}}$ and $I_{\text{a,II}}$, we find:

$$v_{\text{I}} = \frac{I_{\text{a,I}}}{n_{\text{I}} F A_{\text{I}}} = -\frac{i_{\text{l,II}} A_{\text{II}}}{n_{\text{I}} F A_{\text{I}}} \tag{7.12}$$

7.1.2 Galvanic corrosion

Galvanic corrosion is an accelerated corrosion of a metal due to formation of a corrosion cell with a metal or non-metallic conductor that exhibits a higher corrosion potential. For example, if a water pipe made of zinc-coated steel (galvanized steel) is connected to a brass fixture and caution is not taken to electrically isolate the two metals, a corrosion cell is established (Figure 7.5). To simplify the situation, we have replaced, in Figure 7.5, the zinc-coated steel by pure zinc and the brass by copper. The cathodic reaction is the reduction of dissolved oxygen, which takes place on both metals. The corrosion cell formed between the zinc and the copper leads to an accelerated corrosion of zinc near the joint.

In order to better understand this phenomenon, we turn to the model shown in Figure 7.5(b), consisting of a zinc and a copper electrode, both in contact with the same solution. The electrolyte contains neither zinc nor copper ions. In this respect, the described electrochemical cell is very different from the Daniell cell presented in Chapter 2, where copper and zinc electrodes are in equilibrium with their own ions. In contrast, in Figure 7.5 (b), each electrode is in contact with an inert electrolyte (drinking water) and an oxidizing agent (dissolved oxygen). Electrically isolated from each other, each metal assumes its own corrosion potential and corrodes at a given rate but as soon as an electrical contact is established between the two electrodes, a current flows from the copper to the zinc. Due to the resulting polarization, the

© 2007, first edition, EPFL Press

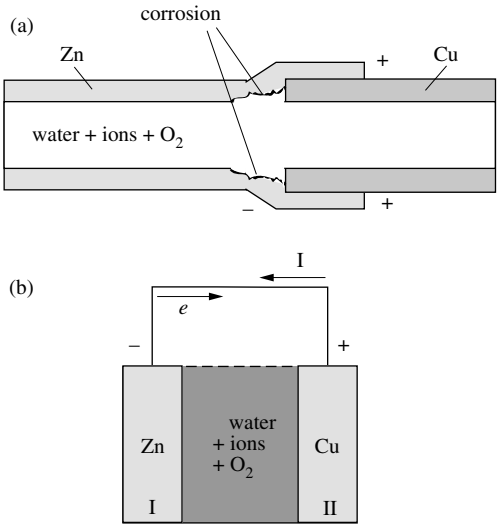


Figure 7.5 (a) Galvanic corrosion between zinc and copper in natural water; and (b) the corresponding electrochemical cell.

potential of the copper electrode shifts in the negative direction and that of the zinc electrode in the positive direction.

The Evans diagram shown in Figure 7.6 schematically describes this situation. Raising the potential from $E_{cor,I}$ to E_I causes an increase of the anodic partial current of the zinc. Simultaneously, the anodic current of the copper decreases as the potential drops from $E_{cor,II}$ to E_{II}.

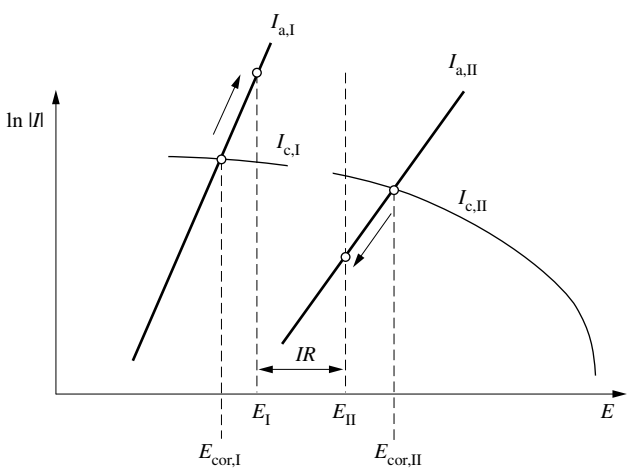


Figure 7.6 Evans diagram of a galvanic corrosion cell showing the increase of the anodic current at the less noble electrode and the reduction of the anodic current at the more noble electrode.

© 2007, first edition, EPFL Press

The potential difference ΔE is equal to the sum of all ohmic resistances of the cell, multiplied by the current ($\Delta E = E_{II} - E_I = IR$). The effect of galvanic coupling on corrosion rate is therefore all the more pronounced as the internal and external resistances are small. Because the ohmic resistance increases with distance, the corrosion rate of the anode decreases with increasing the distance between the two electrodes (*distance effect*). In Figure 7.5(a), we expect the corrosion rate of the zinc to reach its maximum in the immediate vicinity of the copper-zinc contact. Further away, the potential of the zinc gradually approaches the corrosion potential and the corrosion rate decreases to a value one would observe in absence of galvanic coupling. In many drinking water systems, solid reaction products precipitate on the surface, leading to build up of a scale, consisting mostly of calcium carbonate. The porous film thus formed adds an additional internal resistance to the corrosion cell and thus attenuates the effect of galvanic coupling. In the presence of well-formed carbonate scales, galvanized steel piping systems can therefore resist galvanic corrosion for many years.

Galvanic series

The difference in corrosion potential between two metals provides the driving force for galvanic corrosion. Because the corrosion potential is a kinetic quantity, its value depends on a multitude of factors, including certain characteristics of the environment (Chapter 4). A metal, therefore, does not have a single and unique corrosion potential that can be listed in a Table and any compilation of corrosion potentials refers to specific environmental conditions. A list of corrosion potentials of different metals for a given environment is called a ***galvanic series***. Table 7.7 provides a galvanic series established in flowing, aerated sea water.

Table 7.7 Galvanic series in flowing, aerated seawater [1].

Metal	E_{cor} (V)
magnesium	–1.4
zinc	–0.8
aluminum and aluminum alloys	–0.8 to –0.5
carbon steel, cast iron	–0.5 to –0.4
brass	–0.2 to –0.05
304 stainless steel, active	–0.3
316 stainless steel, active	–0.2
lead	–0.3
brass 60-40	–0.26
copper	–0.07
nickel, passive	0.1
304 and 316 stainless steels, passive	0.2
graphite	0.5

© 2007, first edition, EPFL Press

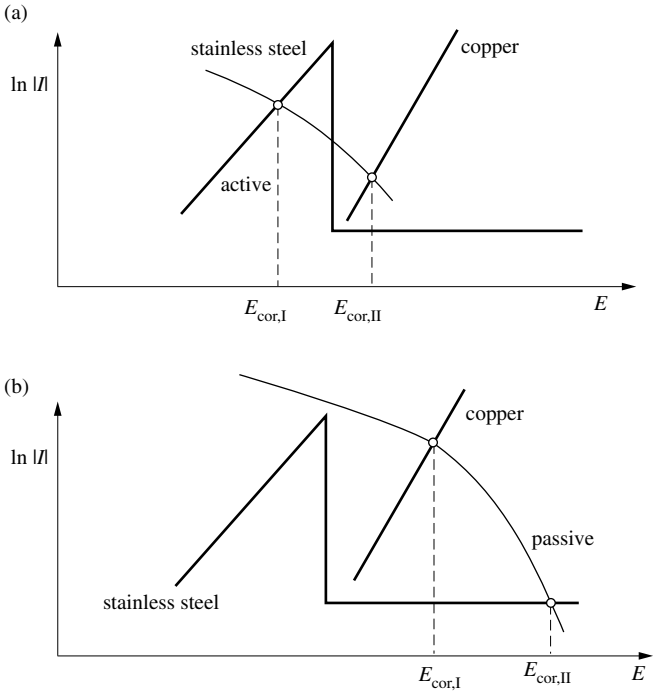


Figure 7.8 Evans diagram showing the galvanic coupling between stainless steel and copper in (a) the active and (b) the passive states.

Galvanic series are fundamentally different from the series of standard potentials of electrode reactions (Chapter 2), which is based entirely on thermodynamic quantities. In a galvanic series, the value of the potential and the relative positions of the metals in the series can vary depending on the environment. Furthermore, individual metals may appear twice in the same galvanic series, once each for the active and the passive states. For example, according to Table 7.7, stainless steel in the active state is less noble than copper, while for the same steel in the passive state

Table 7.9 Galvanic series for different metals in aerated humid soil.

Metal	E_{cor} / V
zinc	–0.7 to –0.4
carbon steel, new	–0.5 to –0.2
lead	–0.4 to –0.2
carbon steel, rusted	–0.3 to –0.1
rusted cast iron	–0.1 to 0.1
copper	0.1 to 0.3

© 2007, first edition, EPFL Press

the situation is reversed. The Evans diagrams of Figure 7.8 further illustrate this behavior.

Table 7.9 gives the corrosion potentials for different metals in humid, aerated soil. The values for new and rusted steel differ because the oxide layers that form during corrosion modify the reaction kinetics, thus increasing the corrosion potential. For example, when a section of a buried pipe is replaced without taking the precaution to electrically isolate the old part from the new one, galvanic corrosion may ensue.

Corrosion rate

The rate of galvanic corrosion depends in a complicated way on the interactions between the two metals. Here we shall estimate the corrosion rate for the two limiting cases already considered in Section 7.1.1:

- ohmic control
- control by oxidant transport

If the ohmic resistance controls the rate of galvanic corrosion, equation (7.11) is applicable. The Evans diagram shown in Figure 7.10 schematically represents this case, when $E_{\mathrm{I}} \approx E_{\mathrm{cor,I}}$ and $E_{\mathrm{II}} \approx E_{\mathrm{cor,II}}$.

On the other hand, if the transport rate of an oxidizing agent is rate limiting, equation (7.12) is valid. In the derivation of this equation, we assumed that the cathodic partial current at the less noble electrode is negligible, which is not always true. Let us assume that oxygen is reduced at both electrodes at the limiting current. In this case, we must take into account the cathodic partial current at each electrode. If the limiting current densities are identical ($i_{\mathrm{c,I}} = i_{\mathrm{c,II}} = i_{\mathrm{l,O2}}$), the corrosion rate of the less noble metal will be proportional to the surface ratio $(A_{\mathrm{I}} + A_{\mathrm{II}})/A_{\mathrm{I}}$.

$$v_{\mathrm{I}} = \frac{i_{\mathrm{a,I}}}{n_{\mathrm{I}}F} = -\frac{i_{\mathrm{l,O_2}}}{n_{\mathrm{I}}F}\,\frac{A_{\mathrm{I}} + A_{\mathrm{II}}}{A_{\mathrm{I}}} \tag{7.13}$$

When the anode has a much smaller surface than the cathode ($A_{\mathrm{I}} << A_{\mathrm{II}}$) this equation reduces to (7.12).

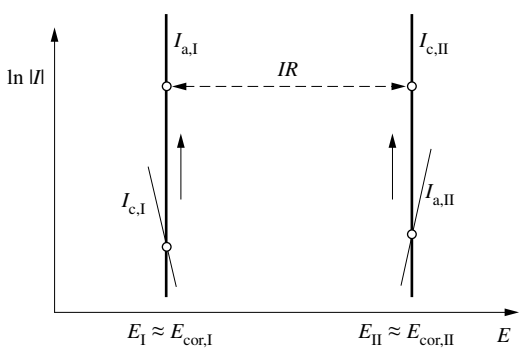


Figure 7.10 Evans diagram for galvanic corrosion under ohmic control.

© 2007, first edition, EPFL Press

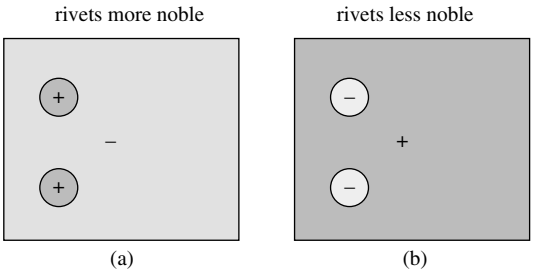


Figure 7.11 Galvanic corrosion of a riveted structure: (a) rivets more noble and (b) rivets less noble than the sheet metal.

A metallic structure secured with rivets or screws (Figure 7.11) offers a particularly good illustration of how the surface ratio between anode and cathode affects the corrosion rate; galvanic corrosion is much more damaging if the rivets, with their small surface area, are anodic relative to the sheet metal than vice versa. For the same reason, a weld requires solder whose corrosion potential is equal to or exceeds that of the structure.

Displacement reactions

Even in the absence of an electrical contact between two parts of a structure made of different metals, galvanic corrosion phenomena can still occur locally, when displacement reactions take place. A well know example are domestic heating systems consisting of a water heater made of copper or brass combined with galvanized steel piping (Figure 7.12). Intrinsic corrosion of the water heater liberates copper ions that

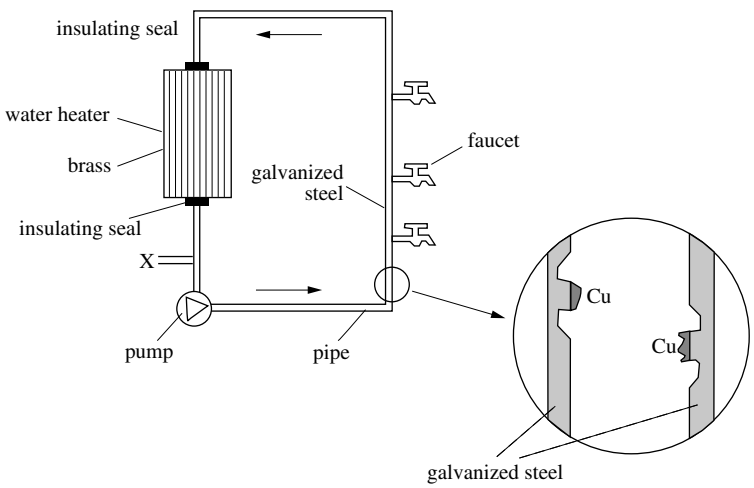


Figure 7.12 Copper deposition on zinc coated steel in a heating system leading to galvanic corrosion.

© 2007, first edition, EPFL Press

are entrained by the hot water into the galvanized steel piping. Copper deposits onto the pipe surface according to the displacement reaction (2.67):

$$Cu^{2+} + Zn \rightarrow Cu + Zn^{2+} \tag{2.67}$$

By itself, this reaction does not pose a great danger because only small amounts of Cu^{2+} are available for reaction. However, due to the displacement reaction nuclei of metallic copper form at the pipe walls and cause galvanic corrosion of zinc and steel in their immediate vicinity. As a consequence, one observes the formation of pits that eventually lead to perforation of the pipe walls.

7.1.3 Aeration cells

A difference in oxygen accessibility between two parts of a metallic structure can induce a corrosion cell, called ***differential aeration cell*** or simply ***aeration cell***. In Figure 7.13, a pipe assembly with a flange is shown. We assume that the pipe is filled with a flowing aerated electrolyte. The flat rubber seal shown is undersized, thus giving rise to a small crevice between the flanges. Inside the crevice, convection is much weaker than outside and as a consequence mass transport of oxygen to the metal surface is slower. An aeration cell is created, which accelerates corrosion in the crevice between the two flanges.

To better understand the described behavior we consider the model of an oxygen concentration cell schematically shown in Figure 7.14. Two identical electrodes, (I) and (II), are in contact with an electrolyte, the same for both electrode compartments. A frit prevents mixing of the two solutions, into which mixtures of nitrogen and oxygen are injected. The oxygen concentration of the gas on the side of electrode (II) is higher than that on the side of electrode (I). Since according to Henry's law, the equilibrium concentration of dissolved oxygen is proportional to its partial pressure in the gas, the oxygen concentration of the electrolyte in compartment (II) is higher than in compartment (I).

If there is no electrical contact, each electrode individually assumes a corrosion potential E_{cor}. In the case of ***non-passivating metals***, both corrosion potentials are

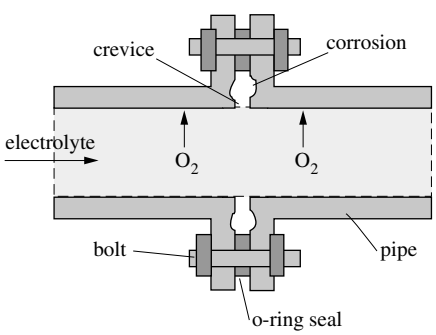


Figure 7.13 Crevice corrosion in a pipe system.

© 2007, first edition, EPFL Press

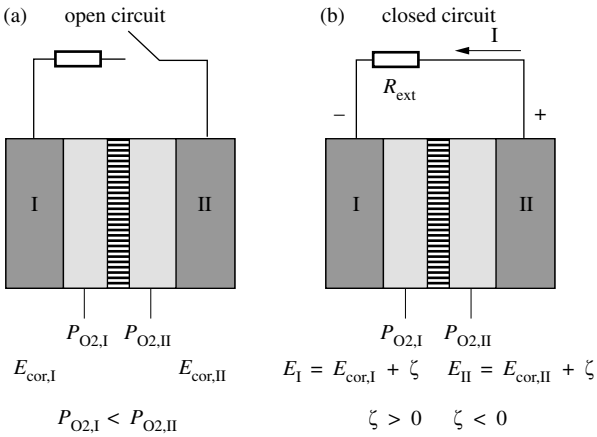


Figure 7.14 Principle of aeration cells operating under (a) open and (b) closed circuit conditions. R_{ext} corresponds to the external resistance.

located in the active region, and according to the Nernst equation (7.15), applied to the reaction $O_2 + 4H^+ + 4e = 2\ H_2O$, the corrosion potential of electrode (II) is the more noble of the two:

$$E_{rev} = E^\circ + \frac{RT}{4F} \ln p_{O_2} a_{H^+}^4 \tag{7.14}$$

The actual difference in corrosion potential $\Delta E_{cor} = E_{cor,II} - E_{cor,I}$ depends on the anodic kinetics and can not be predicted solely from equation (7.15). This is illustrated by Fig. 7.15(a) where the oxygen reduction is controlled by mass transport. Once an electrical contact is established between the two electrodes (Figure 7.15(b)), the potential of electrode (I) increases and that of electrode (II) decreases. This accelerates the corrosion of electrode (I), exposed to the solution of lower oxygen concentration and slows down the corrosion of the electrode (II) exposed to the solution of higher oxygen concentration. It also follows from the Figure that the corrosion rate of electrode (I) may reach but never exceed that of electrode (II). This conclusion is valid as long as we deal with non-passivating metals.

In case of ***passivating metals*** differential aeration cells can have a much more striking effect than observed with active metals. Figure 7.16 shows two Evans diagrams analogous to those of Figure 7.15, but for a passivating metal. In Figure 7.16(a) electrode (II) is passive while electrode (I) is active. This situation can occur if the oxygen concentration in compartment (I) is not sufficient for spontaneous passivation (Chap. 6). When electrical contact is established between the electrodes, the corrosion of electrode (I) is accelerated while that of electrode (II) does not markedly change (Figure 7.16(b)). For small passive currents ($I_{a,II} \approx 0$), we get from equation (7.4) :

© 2007, first edition, EPFL Press

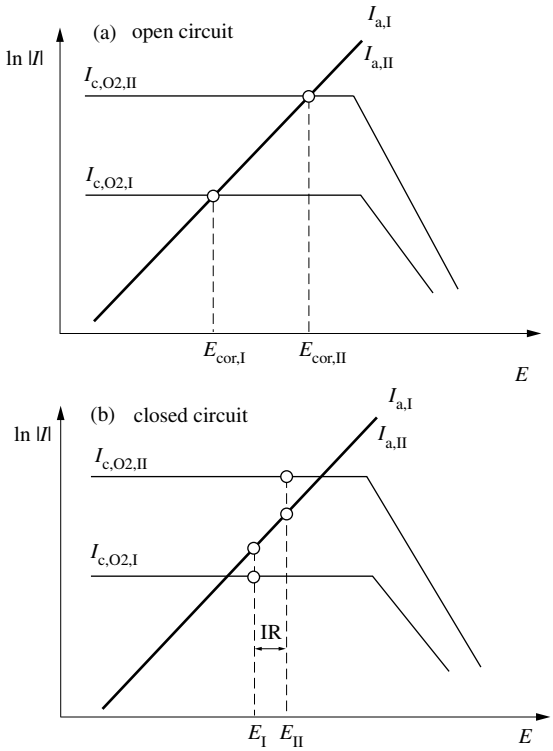


Figure 7.15 Evans diagram for an aeration cell involving an active metal under (a) open circuit and (b) closed circuit conditions. *IR* represents the sum of the ohmic drops in the system.

$$i_{a,\,I} \approx -\frac{I_{c,\,I} + I_{c,\,II}}{A_I} \tag{7.15}$$

When the surface area of electrode (I) is small compared to that of electrode (II) and/or the oxygen concentration in compartment (I) is much lower than in compartment (II), the cathodic partial current of electrode (I) can be neglected: $|I_{c,I}| << |I_{c,II}|$. By setting $I_{c,II} = i_{c,II}\, A_{II}$, we obtain thus for the anodic partial current density of electrode (I):

$$i_{a,\,I} = -\, i_{c,\,II}\, \frac{A_{II}}{A_I} \tag{7.16}$$

Equation (7.16) predicts that in a differential aeration cell consisting of an active and a passive electrode (***active-passive cell***), the corrosion rate is proportional to the surface ratio A_{II}/A_I between cathode and anode. A large surface area ratio strongly enhances the corrosion rate of the anode. When active-passive cells are formed, the corrosion rate of those parts of a structure that are less accessible to oxygen can vastly exceed that of the parts that are more accessible!

© 2007, first edition, EPFL Press

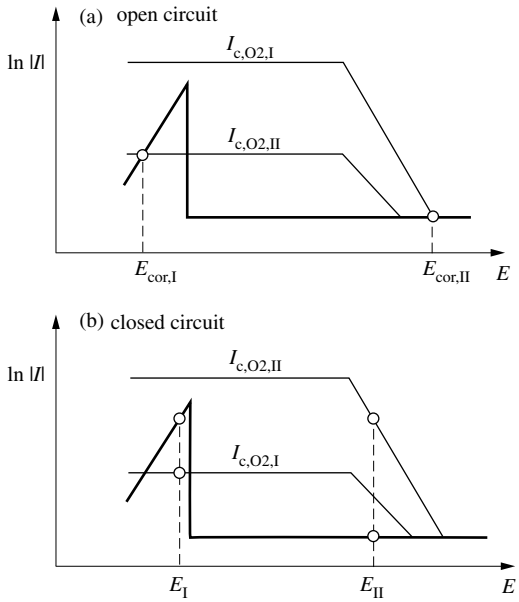


Figure 7.16 Evans diagram of oxygen cells involving passivating metals under (a) open circuit and (b) closed circuit conditions.

In Figure 7.13 the pipe has a much larger surface area than the crevice. If the pipe and the flanges are made of stainless steel, a passivating metal, the surface in the crevice can become active by lack of oxygen, while the pipe wall remains passive. As a consequence the corrosion rate in the crevice is greatly accelerated. This type of localized attack is known as ***crevice corrosion.*** Crevice corrosion is one of the most common type of localized corrosion. It is usually the result of a differential aeration cell involving a passivating metal that becomes locally active. Generally, the occurrence of crevice corrosion is enhanced by the presence of chloride ions, which facilitate establishment of the active state in the crevice.

Effect of the local pH at the metal surface

The establishment of active-passive cells is facilitated by pH changes induced by the anodic and cathodic partial reactions. To understand this behavior, we recall first that passivation of metals is favored by a high pH (Chapter 6). The Evans diagram of Figure 7.17 shows the anodic partial current of a passivating metal exposed to an aerated electrolyte. Curve (b), which exhibits a lower passivation current than curve (a), corresponds to a solution of a higher pH. Oxygen is assumed to be reduced at the same rate on both metals. Clearly, at open circuit the electrode (a) is in the active state, while the electrode (b) is in the passive state. Electrical contact between the two therefore leads to the establishment of a corrosion cell in which metal (b) is the cathode.

© 2007, first edition, EPFL Press

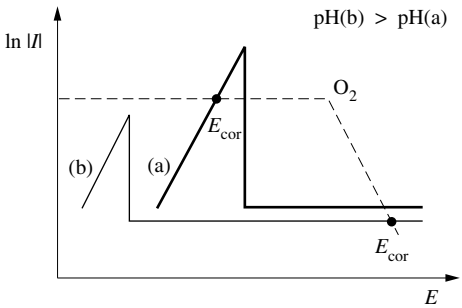


Figure 7.17 Evans diagram showing the effect of pH on corrosion potential of passivating metals in the presence of oxygen.

The reduction of oxygen under acidic conditions leads to a consumption of protons:

$$O_2 + 4\,H^+ + 4\,e \rightarrow 2\,H_2O \tag{7.17}$$

while under neutral or alkaline conditions it produces hydroxyl ions:

$$O_2 + 2\,H_2O + 4\,e \rightarrow 4\,OH^- \tag{7.18}$$

In both cases, the local pH at the surface increases, the extent of the increase depending on the rate of mass transport of the protons or hydroxyl ions. Similarly, cathodic hydrogen formation by the reduction of protons also leads to a local increase in pH.

On the other hand, anodic reactions tend to lower the local pH near the electrode. The dissolution of iron, for example, produces protons by the hydrolysis equilibrium of ferrous ions:

$$Fe^{2+} + H_2O = FeOH^+ + H^+ \tag{7.19}$$

Other metals, including nickel, chromium and cobalt, react in a similar fashion.

Generally, under conditions of uniform corrosion, the effect on pH of the anodic and cathodic partial reactions cancel each other. In contrast, under conditions of localized corrosion, the anodic and cathodic partial reactions are locally separated and a pH difference builds up between anodic and cathodic sites.

Evans demonstrated this behavior with a very simple experiment. A drop of water containing dissolved ions, along with a pH indicator, is placed on a plate of carbon steel and exposed to the atmosphere. The surface at the center of the droplet starts to corrode, whereas the surface near the edge remains intact. The color of the indicator turns orange at the center, thus denoting an acidic pH, and blue towards the exterior, indicating a basic pH. The presence of the water drop on the steel surface permits the formation of an aeration cell because oxygen has limited accessibility to the drop's

© 2007, first edition, EPFL Press

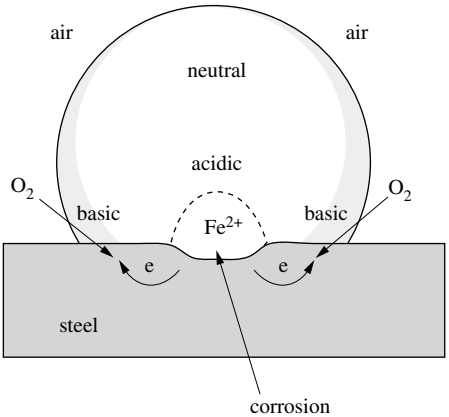


Figure 7.18 Evans experiment showing pH change within a water droplet due to different accessibility of oxygen.

center (Figure 7.18). The center thus becomes anodic with respect to the exterior boundary and the local pH decreases due to formation of ferrous ions:

$$Fe \rightarrow Fe^{2+} + 2\,e \tag{7.20}$$

The ferrous ions can hydrolyze according to (7.19) or oxidize to ferric ions, which also exhibit hydrolysis equilibra that lower the pH.

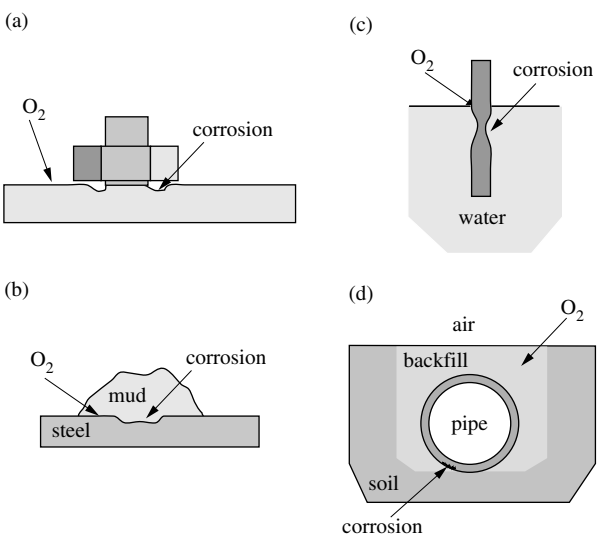


Figure 7.19 Examples of corrosion due to aeration cells.

© 2007, first edition, EPFL Press

At the exterior boundary of the drop, the pH rises because the cathodic partial reaction dominates:

$$\frac{1}{2}\,O_2 + H_2O + 2\,e \rightarrow 2\,OH^- \tag{7.21}$$

The result is an accelerated corrosion at the center of the drop, where the steel eventually becomes active, while the surface at the edge remains passive.

Differential aeration cells are frequently encountered in practice. They form in particular on carbon steel in neutral environments, whenever oxygen accesses certain parts of the structure with more difficulty. Figure 7.19 gives a few examples: (a) presence of a crevice between a poorly tightened nut and a plate; (b) accumulation of muddy residue on a steel plate; (c) a steel structure partially immersed in water; (d) a buried pipe that is not entirely surrounded by aerated backfill.

7.2 NON-UNIFORM CORROSION GOVERNED BY METALLURGY

7.2.1 Effect of microstructure on local dissolution rate

During active metal dissolution atoms are removed from the crystal lattice and pass into solution as hydrated or complexed ions. This process does not occur statistically all over the surface, but only at energetically favorable sites. On low index surfaces, dissolution takes place from kink sites located at mono-atomic steps (Chapter 3). Depending on prevailing electrochemical conditions, two reaction paths are then possible (Figure 7.20):

- the atom ionized at the kink site passes directly into solution in hydrated form (a);
- the atom moves from the kink site by surface diffusion forming an adsorbed atom (adatom) that subsequently undergoes ionization and hydration (b).

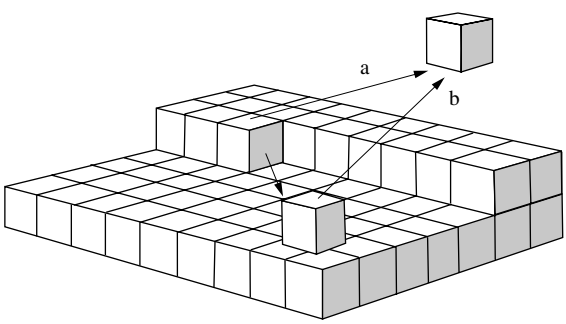


Figure 7.20 Dissolution of metal atoms from kink sites by (a) direct transfer into solution; and (b) intermediate formation of adatom.

© 2007, first edition, EPFL Press

For both reaction paths, as atoms leave from kink sites, the mono-atomic steps are displaced on the surface. If the number of kink sites on a step does not depend on the distance between steps, the dissolution current density is equal to [2]:

$$i_a = v_{st}(E) \cdot \rho_{st} \cdot q_{mon} \tag{7.22}$$

where $v_{st(E)}$ designates the average velocity of the atomic steps (cm/s), which depends on potential. The atomic step density ρ_{st} is the inverse of the average distance between two steps ($\rho_{st} = 1/L_{st}$), and q_{mon} is the charge required to dissolve a monolayer (C/cm^2).

According to (7.22), the rate of anodic dissolution at a given potential increases with the density of atomic steps. For slightly misoriented planes with respect to a low index plane, it is therefore expected to vary with the angle of misorientation.

Crystalline attack

Real surfaces contain many defects such as vacancies, dislocations, grain boundaries, inclusions, and segregations. Furthermore, they are not atomically flat, but have a rugosity that depends on the pretreatment. Equation (7.22) therefore only represents an incomplete view of the phenomena taking place on an atomic scale during dissolution. However, the simple model allows us to understand certain effects, such as the orientation dependent ***formation of facets***, which is important for the metallographic study of grain structure of polycrystals by chemical or electrochemical etching. The adsorption of anions or organic molecules at a kink site can block dissolution and thus slow down the propagation of a mono-atomic step (Figure 7.21). Another step, whose movement is not affected by adsorbed species, can catch up to the blocked step and giving rise to the formation of a step double in height. If this process repeats itself, multi-atom steps can form that are of sufficient height to be visible with a scanning electron microscope or even an optical microscope. The geometry of the facets formed reflects the crystalline orientation of the grains, and thus allows us to study the microstructure of a polycrystalline material (Figure 3.34).

Defects that intersect the surface perturb the crystalline structure and therefore facilitate dissolution. For example in Figure 3.34 the grain boundaries have been preferentially attacked. Grain boundaries are locations where the crystal structure is strongly perturbed and the atoms therefore are less solidly bound. Inclusions are other

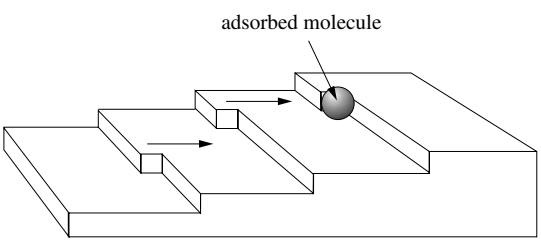


Figure 7.21 Formation of multiatom steps by blocking due to adsorption.

© 2007, first edition, EPFL Press

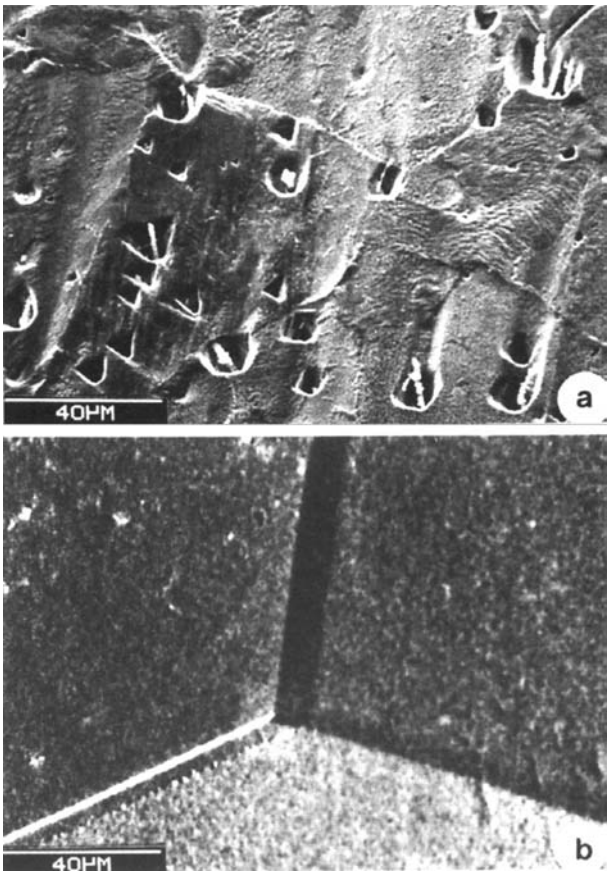


Figure 7.22 SEM pictures of nickel surfaces after anodic dissolution: (a) commercial grade nickel (Ni 200) ; and (b) high purity nickel [3].

defects favoring dissolution. Figure 7.22 shows the surface of commercial-grade nickel (a) and of high purity nickel (b) after having been subjected to identical anodic dissolution conditions in a chloride solution. In the first case, the dissolution reveals the presence of a large number of pits having formed at sulfide inclusions which act as preferred dissolution sites. No such pits are visible on the high purity nickel that does not contain inclusions, although the dissolution still reveals the grain structure.

Corrosion behavior of amorphous alloys

By rapid solidification, it is possible to produce alloys that show an X-ray amorphous structure. Typically, these alloys contain a large fraction of non-metallic elements, such as boron or phosphorus, which slow down crystallization. Atoms in amorphous alloys are less densely packed than in crystalline alloys. Furthermore, amorphous structures are thermodynamically unstable. Both factors should induce a

© 2007, first edition, EPFL Press

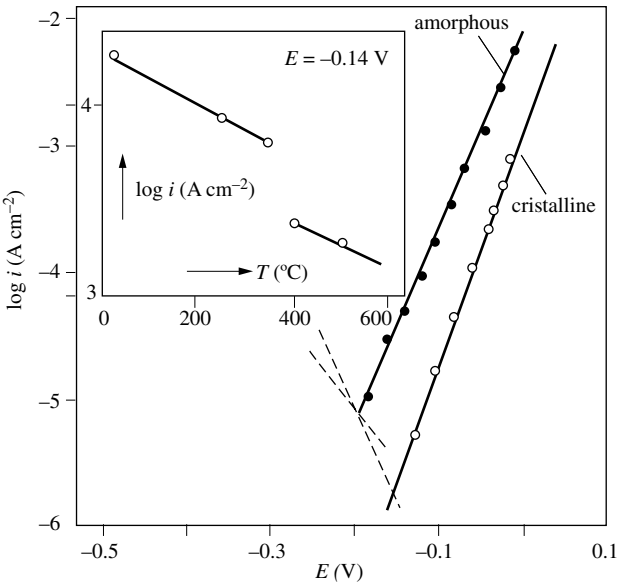


Figure 7.23 Anodic polarization curve of Ni_2B in 0.5M H_2SO_4 and 25 °C in both the cristalline state (•) and the amorphous state (o) [4].

higher dissolution rate compared to a crystalline material of same composition. This has indeed been observed for the nickel boron alloy of composition Ni_2B, which can be made in crystalline or amorphous form [4].

Figure 7.23 shows the anodic polarization curves of Ni_2B in its amorphous state and after a thermal treatment (500 °C, 5h) leading to crystallization. At equal potential, the anodic current density of the amorphous alloy is larger than that of the crystalline material, indicating that the amorphous material dissolves more easily. The insert shows that for a given potential (–0.14 V), the anodic current density as a function of thermal treatment temperature exhibits a discontinuity at 400 °C , which is due to the phase transition from amorphous to crystalline.

Some amorphous alloys are highly corrosion resistant. The reason is their unusual chemical composition and the absence of phase boundaries. Due to the high solidification rates applied, there is not enough time for segregation of elements and crystallization of distinct phases. The alloy composition is therefore not constrained by solubility limits and one can add higher concentrations of corrosion resisting alloying elements than in crystalline materials. For example, certain amorphous alloys based on iron, chromium and nickel are extremely corrosion resistant because they contain large amounts of chromium and molybdenum, in addition to non-metallic elements such as phosphorus and boron. However, these materials can only be produced economically as thin sheets or coatings and they are often brittle, which limits their practical use.

© 2007, first edition, EPFL Press

Multiphase alloys

The corrosion behavior of multiphase alloys depends on the electrochemical properties of the individual phases present and on their galvanic interactions. Similar as in galvanic corrosion, we expect the dissolution of a less noble phase to be accelerated by the presence of a nobler phase, the importance of the effect increasing with the difference in corrosion potentials. An example is presented in Figure 7.24 [5]. The scanning electron micrograph (a) shows the surface of a duplex stainless steel after chemical etching in a concentrated HCl solution. The picture reveals a two phase structure (ferrite and austenite), the ferrite phase being attacked preferentially. The image (b) was taken after exposing the previously etched surface to a copper sulfate solution. The ensuing displacement reaction between the steel and the copper sulfate leads to precipitation of copper. The SEM image shows that copper deposits selectively on the austenitic phase, which is more cathodic than the ferrite phase.

The ***graphitic corrosion*** of gray cast iron is a particularly severe case of galvanic corrosion of a two phase material. Gray cast iron contains graphite embedded in a ferrite phase. The graphite being electrochemically more noble than the ferrite, a selective attack of the latter is observed under certain corrosion conditions, rendering the material porous and flaky. Similar galvanic effect can, in principle, take place in any multiphase material. ***Metal matrix composites*** typically contain an aluminum

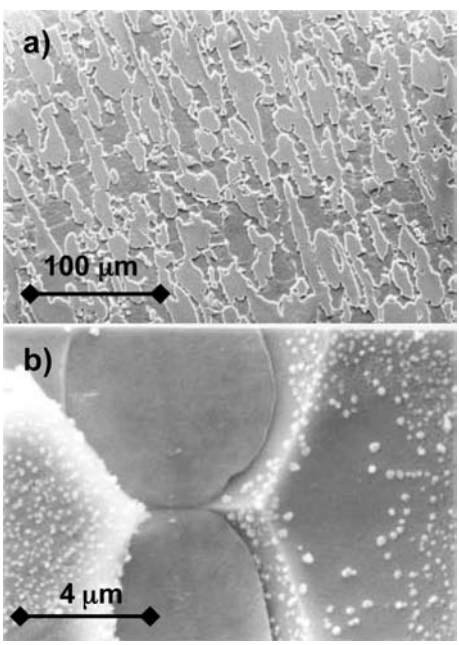


Figure 7.24 Scanning electron micrographs of the surface of a duplex stainless steel after immersion in 10% HCl for 2 h showing selective attack of constituent phases (a). The micrograph (b) shows the same alloy after immersion in 0.1 M $CuSO_4$. One observes selective deposition of copper on the more cathodic austenite phase [5].

© 2007, first edition, EPFL Press

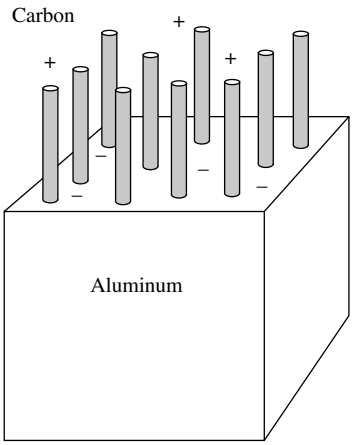


Figure 7.25 Corrosion cell formed between graphite fibers and an aluminum matrix (schematic).

matrix reinforced by high strength ceramic fibers such as boron or carbon. The latter are electronically conducting and therefore, unless they are protected by an insulating coating, they may form the cathode of a galvanic cell with the aluminum matrix as the anode (Figure 7.25).

Because corrosion resistant engineering alloys are generally passive, differences in corrosion potential between individual phases are mostly small. Indeed, the example of Figure 7.24 refers to chemical etching in the active state. In the passive state the differences between phases would be less pronounced because they are attenuated by the properties of the passive film. Galvanic interactions between alloy phases therefore are generally not a major problem in corrosion engineering, except for grain boundary

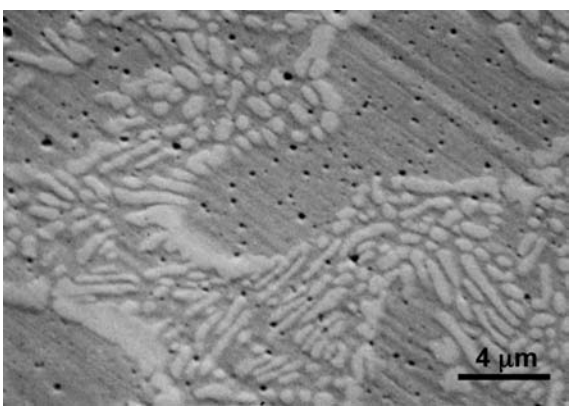


Figure 7.26 Scanning electron micrograph of pitted surface of an experimental Fe-25Cr-11Nb alloy exhibiting a niobium poor phase (dark) and a niobium rich phase (bright). Pits form preferentially on the dark phase [6].

© 2007, first edition, EPFL Press

precipitates and inclusions that will be discussed later in this chapter. On the other hand, the pitting resistance of different phases in a passive alloy may differ. The micrograph of Figure 7.26 shows the microstructure of an experimental Fe-Cr-Nb alloy with an interdendritic phase (bright color) rich in niobium and a dendritic phase containing less niobium. During anodic polarization in a chloride medium, pits are observed only in the niobium poor phase.

In metallography, differences in electrochemical properties of alloy phases are commonly used to visualize the microstructure of multi-phase alloys. The strongest selectivity is observed for conditions (electrolyte, potential) where one phase is active and the other passive.

7.2.2 Dealloying

The term ***dealloying*** designates a corrosion process that involves the selective dissolution of a constituent element from an alloy phase. For example, during dealloying of brass (*selective corrosion of brass*) the zinc is leached selectively from the alloy, leaving a flaky and porous layer of copper. The process is therefore also known as ***dezincification***. The vulnerability of brass with respect to dealloying increases with the zinc content: yellow brass (30% Zn, 70% Cu) and Muntz metal (40% Zn, 60% Cu) are quite sensitive to this type of corrosion, while red brass (15% Zn, 85% Cu) is much less so. The addition of 1% of tin, or of such elements as As, Sb, P (admiralty brass) greatly improves the resistance of yellow brass to dealloying. Certain gold-copper alloys exhibit selective corrosion of copper. In humid environments containing sulfides this leads to formation of a brownish-black tarnish layer.

Dealloying can also serve a useful purpose. An well-known example is depletion-guilding, a process that has been used since ancient times for enriching gold alloy sheets with gold. A more recent application is the fabrication of nano-porous metal by selective leaching of the less noble element from a single-phase alloy. Figure 7.27

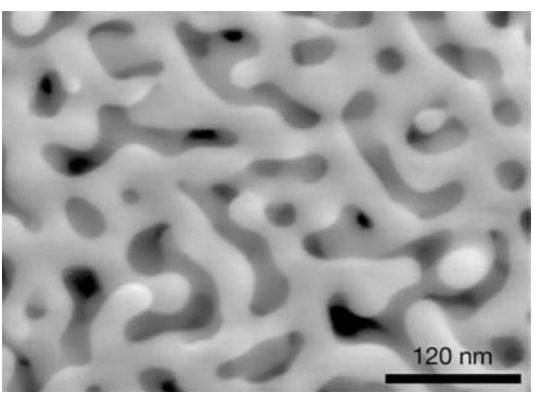


Figure 7.27 Scanning electron micrograph of a dealloyed gold-silver alloy. Immersion in concentrated nitric acid led to selective dissolution of silver from the alloy leaving a gold rich nanoporous structure [7].

© 2007, first edition, EPFL Press

shows the nano-porous structure of gold formed by exposing a silver-gold alloy to nitric acid, which selectively corrodes the silver.

Polarization behavior of binary single-phase alloys

Binary alloys whose phase diagram shows solid solubility over a wide concentration range are particularly well suited for the study of dealloying phenomena. Figure 7.28 schematically presents the pseudo-steady state anodic polarization curve of a binary single-phase alloy AB and of its two constituents, the element B being more noble than A. Two distinct potential regions are observed, separated by the ***critical potential***, E_{crit}:

- the subcritical region, $E < E_{crit}$;
- the overcritical region, $E > E_{crit}$.

In the subcritical region, the anodic partial current density varies little with potential. Its value, generally small, depends on the applied potential sweep rate and on the alloy

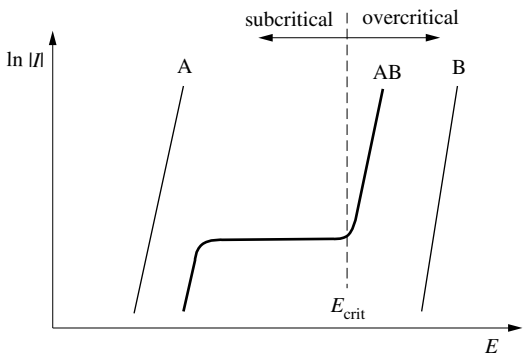


Figure 7.28 Polarization curve of binary single phase alloy (schematic).

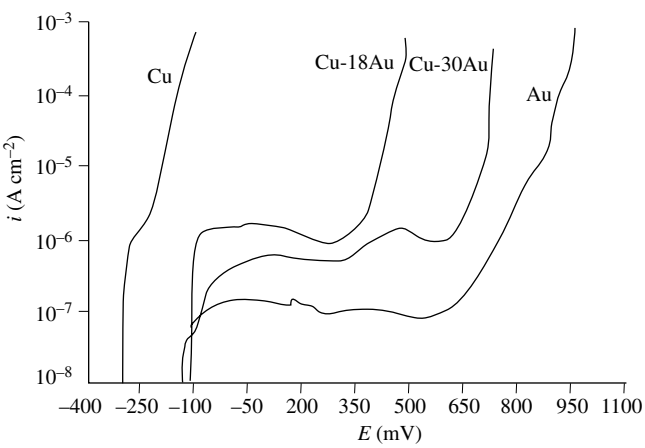


Figure 7.29 Anodic polarization curves of copper, gold and two Cu-Au alloys in 12 M LiCl, 25 °C. Sweep rate is 0.1 mV s^{-1} [8].

© 2007, first edition, EPFL Press

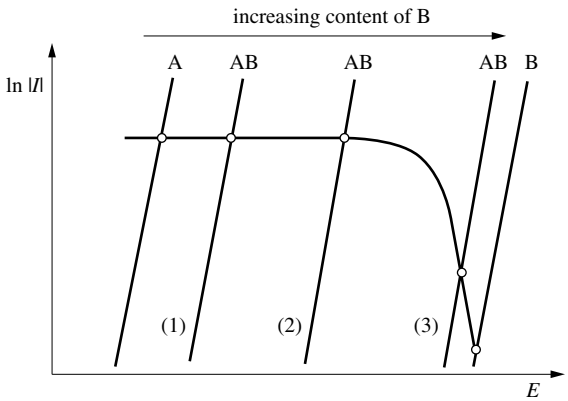


Figure 7.30 Evans diagram showing the effect of alloy composition on corrosion potential in aerated solution.

composition. In contrast, in the overcritical region, the anodic current strongly increases with potential.

The behavior is illustrated by the anodic polarization curves of two gold-copper alloys and the pure metals shown in Figure 7.29, which were measured in a concentrated chloride electrolyte capable of dissolving copper as well as gold. The value of the critical potential increases with the noble-metal content in the alloy. The current in the subcritical potential region is lower at higher gold concentration.

The Evans diagram of Figure 7.30 shows how the anodic polarization behavior of an alloy AB affects its corrosion rate in presence of an oxidizing agent such as dissolved oxygen. It is assumed that mass transport controls the rate of oxygen reduction over a wide potential range. A comparison with Figure 7.28 suggests that to resist corrosion, the corrosion potential of the alloy must lie below the critical potential ($E_{cor} < E_{crit}$).

Subcritical potential region

The anodic current in the subcritical region results from the selective dissolution of the less noble metal, in this case A, according to the following reaction:

$$AB \rightarrow A^{n+} + B_{(AB)} + n\,e \tag{7.23}$$

Here, $B_{(AB)}$ represents the non-dissolved component B in the alloy. The selective dissolution of A causes an enrichment of B at the surface, which gradually slows the rate of dissolution of A. In principle, this process should stop entirely once the surface of the alloy contains only atoms of the more noble metal (Figure 7.31). If the atoms in the crystal lattice were to remain in their position, the anodic charge needed to stop dissolution of less noble metal would depend only on the alloy composition. In reality, the surface atoms rearrange and the anodic charge needed for dissolution to decrease

© 2007, first edition, EPFL Press

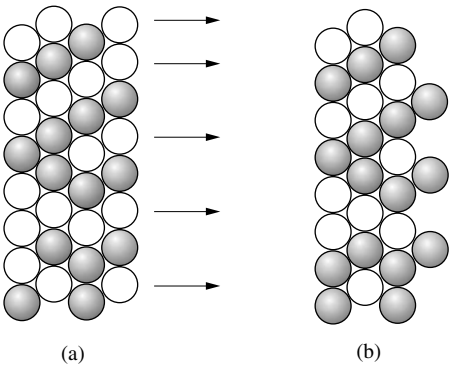


Figure 7.31 Selective dissolution of component A of the binary alloy AB at subcritical potentials: (a) the composition of the surface before dissolution; and (b) the surface enriched in B.

to negligible levels is always higher than that estimated from simple geometrical considerations.

In principle three mechanisms can contribute to rearrangement of the surface following removal of less noble metal atoms (Figure 7.32):

- volume diffusion;
- surface diffusion;
- dissolution with subsequent redeposition of B.

Volume diffusion at room temperature is usually a slow process, but removal of the less noble metal atoms from the crystal lattice could possibly create a high vacancy concentration near the surface and thus lead to a higher diffusion rate than predicted from thermodynamic considerations. Injection of divacancies has also been considered to explain abnormally high solid-state diffusion rates [10]. Surface diffusion is generally faster than bulk diffusion and atoms should be able to rearrange more easily

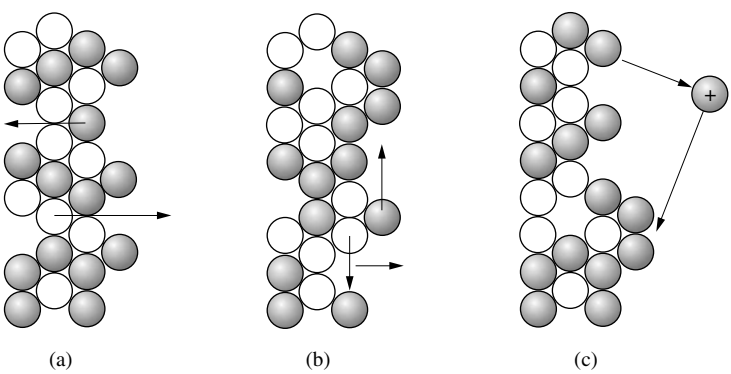


Figure 7.32 Rearrangement of surface atoms of anodically polarized binary alloy AB: (a) volume diffusion; (b) surface diffusion; (c) and dissolution with subsequent redeposition.

© 2007, first edition, EPFL Press

that way [7]. An interesting theoretical model combining microscopic and macroscopic aspects has recently been proposed [7, 11]. By assuming that surface diffusion is the only rearrangement process and the anodic dissolution rate of the less noble component follows Butler-Volmer kinetics the authors were able to numerically simulate the shape of the potentiodynamic anodic polarization curve of a silver-gold alloy, including the variation of the critical potential with alloy concentration. Dissolution of the more noble metal from the alloy followed by redeposition, the third mechanism, is highly improbable at the potentials prevailing in the subcritical region and can be eliminated as a plausible mechanism.

The rate of surface enrichment in the more noble metal can be studied using potentiostatic experiments by measuring the decrease of the anodic current density with time. Figure 7.33 shows results obtained with a mechanically polished Ag-Pd alloy. Also shown is the average atom fraction of palladium at the surface measured by Auger electron spectroscopy, using different polarization times as indicated in the Figure. The data confirm the increasing surface enrichment in Pd.

The observed current decrease with time typical for potentiostatic experiments is not in contradiction with the fact that potentiodynamic experiments show a current plateau in the subcritical potential region. The plateau can be explained by two opposing effects on current. On one hand, surface enrichment diminishes the current with time and hence during a slow forward sweep with increasing potential. On the other hand, if the rate of anodic dissolution of the less noble component is charge transfer controlled, it increases with potential. As a consequence of the two opposing effects an anodic potential sweep leads to a dissolution current that varies little with

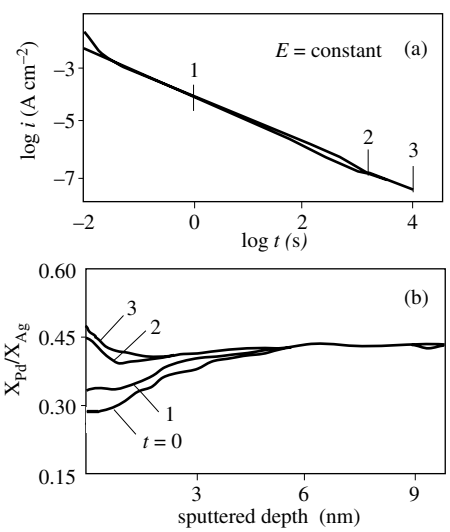


Figure 7.33 Potentiostatic dissolution of Ag-30Pd alloy in 12 M LiCl at constant subcritical potential: (a) the variation of the current with time; and (b) the AES depth profiles showing the Pd/Ag atomic ratio for different polarization times [9].

© 2007, first edition, EPFL Press

potential, but depends on sweep rate and on the concentration of less noble element in the alloy.

Selectivity at overcritical potentials

In the overcritical potential region, anodic dissolution can continue indefinitely, meaning that the anodic reaction is no longer blocked by enrichment of the surface in the more noble component. Depending on the composition of the alloy and the electrochemical properties of its components, one of two types of reactions takes place:

- selective dissolution of the less noble component leaving behind a porous structure of mostly noble metal;
- simultaneous dissolution of both components in the stoichiometric proportion of the alloy.

The ***parting limit*** indicates the minimum concentration of noble element in an alloy above which simultaneous dissolution of both components rather than selective dissolution of the less noble component occurs.

To characterize the transition from selective to simultaneous dissolution we can define the ***selectivity*** S_B, of the dissolution process by equation (7.24):

$$S_B = 1 - \frac{Q_B}{Q_{B,\,th}} \tag{7.24}$$

Here Q_B represents the measured charge, calculated using Faraday's law, that corresponds to the amount of B dissolved. The theoretical charge $Q_{B,th}$, is the charge needed for the dissolution of B when the two elements A and B dissolve according to their stoichiometric proportion in the alloy. For simultaneous dissolution: $Q_B = Q_{B,th}$, $S_B = 0$ and for selective dissolution of A: $Q_B = 0$, $S_B = 1$. In principle, at noble metal concentrations above the parting limit the selectivity is zero, below it is one. In reality the change in selectivity is not abrupt, but occurs over a small concentration interval. This is illustrated by Figure 7.34(a) for three single-phase palladium alloys polarized at constant current ($i = 10^{-4}$ A cm^{-2}) in a concentrated chloride solution. Palladium is the more noble element in the three alloys shown. At a palladium concentration below 10 atom percent ($X_{Pd} \leq 0.1$) only the less noble metal dissolves and $S_{Pd} = 1$. The selectivity drops to zero at $X_{Pd} \geq 0.2$ corresponding to simultaneous dissolution of both alloy elements. Fig. 7.34(b) shows that with the change from selective to simultaneous dissolution, the potential increases from that typical for dissolution of the less noble to that of the more noble element.

When there is a significant difference in reversible potential of the two alloy components, selective dissolution of the less noble element instead of simultaneous dissolution of both elements takes place at overcritical potentials. The same behavior is also observed when one of the elements is in the passive state and the other in the active state or when the more noble alloy constituent is present in only small concentration.

© 2007, first edition, EPFL Press

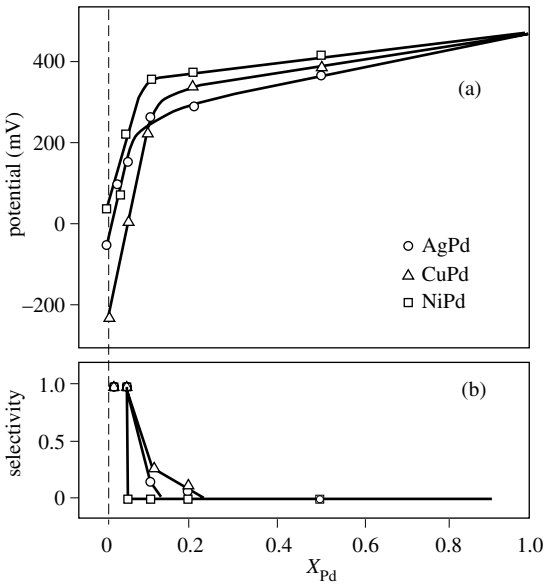


Figure 7.34 Galvanostatic dissolution of Pd alloys in 12 M LiCl, $i = 10^{-4}$ A cm^{-2}: (a) variation of the potenial as a function of Pd content and (b) selectivity coefficients [9].

Gold-copper alloys provide a good example. In concentrated chloride solution, both gold and copper can dissolve, because the reversible potential of gold is lowered by complex formation with chloride ion (Chapter 2). Therefore, while at low gold concentration selective dissolution of copper occurs, both components dissolve from a gold-copper alloy with high gold concentration . The behavior thus corresponds to that of the palladium alloys shown in Figure 7.34. In sulfate solution, on the other hand, gold-copper alloys do not exhibit simultaneous dissolution. The reason is that in the absence of complex formation, the standard potential for dissolution of gold (1.82 V for Au/Au$^+$) is higher than that for the decomposition of water (1.23 V). By anodically polarizing a gold alloy with a concentration exceeding the parting limit, oxygen evolution rather than gold dissolution therefore occurs at high potentials.

Selective dissolution in the overcritical potential region

In the overcritical potential region, contrary to the situation in the subcritical region, selective dissolution of the less noble alloy component does not stop after the removal of a few atomic layers, but continues indefinitely [13]. As dissolution progresses it leaves behind a porous structure of the more noble metal that grows inward from the surface (Figure 7.35). It can reach a thickness of some millimeter, several orders of magnitude more than the enriched surface layers formed at subcritical potentials, which are of nanometer depth only. In practice, corrosion damage due to porous layer formation as a result of selective corrosion results in a lowering of the mechanical resistance of an alloy.

© 2007, first edition, EPFL Press

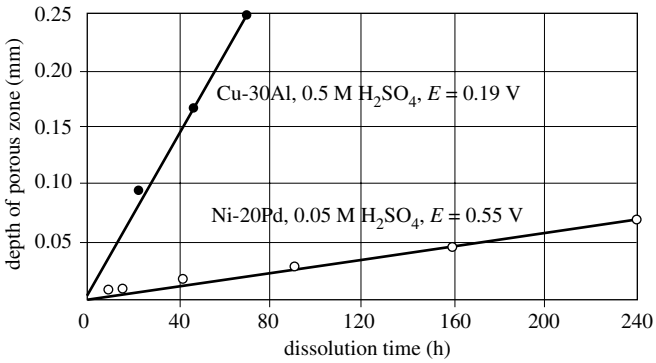


Figure 7.35 Linear growth of porous zone during potentiostatic dissolution of Cu-30 Al (in 0.5 M H_2SO_4) and Ni-20 Pd (in 0.05 M H_2SO_4) at overcritical potential [14].

Three possible mechanisms can explain the formation of porosity above the critical potential. They are schematically presented in Figure 7.36.

If volume diffusion of the less noble metal from the bulk to the surface limits the rate of dissolution**,** a planar surface is not stable, because small waviness is accentuated by local differences in diffusion rate (a). A volume diffusion mechanism, however, cannot explain in a satisfactory way the existence of a critical potential and in addition, solid state diffusion rates are likely to be too slow to account for the observed current densities.

Surface diffusion is a more plausible mechanism for explaining the formation of porosity. Dissolution of the less noble metal liberates atoms of the more noble metal in the form of adatoms. These diffuse on the surface and form aggregates. During this process new atoms of the less-noble metal become exposed to the solution and dissolve, leading eventually to formation of a porous structure of noble metal atoms (b). This model offers an explanation for the origin of the critical potential: the activation energy for diffusion depends only on the number of nearest neighbors, while the dissolution rate of the less noble metal, according to the Butler-Volmer equation, strongly increases with potential. In the subcritical potential region atoms dissolve only from favorable lattice positions such as kinks or steps. Under these

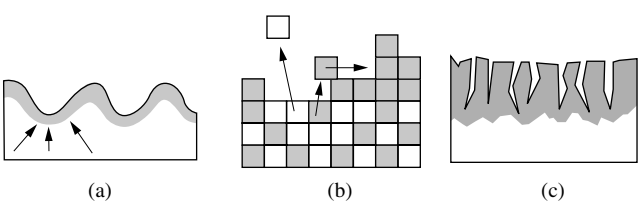


Figure 7.36 Mechanisms of porous layer formation, enriched in noble metal, due to selective dissolution in the overcritical region: (a) instability of a planar surface under dissolution conditions controlled by volume diffusion; (b) atomic rearrangement by surface diffusion; and (c) cracking due to internal stress caused by the coalescence of vacancies.

© 2007, first edition, EPFL Press

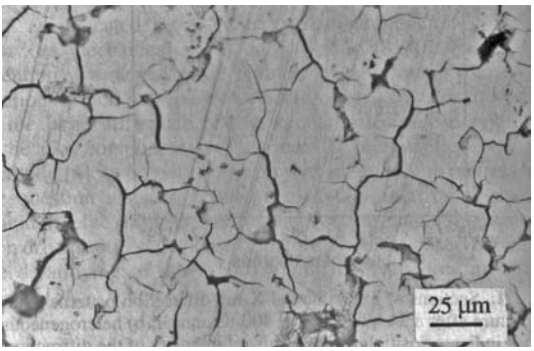


Figure 7.37 Optical micrograph of the surface of a Zn-6.9Fe alloy (delta-phase) after anodic polarization in acidified sodium sulfate solution. Dealloying leads to formation of surface cracks [15].

conditions, atomic rearrangement is able to block the active dissolution sites after removal of just a few atomic layers. At higher potential the less noble metal can dissolve also from less favorable lattice positions such as terraces between the atomic steps. The noble metal atoms left behind form clusters that eventually give rise to a nano-porous structure of the type shown in Figure 7.27. The described model can explain the shape of the anodic current voltage curve and the occurrence of a critical potential provided surface diffusion coefficients some orders of magnitude higher than expected from thermodynamics are assumed [7,11].

A third explanation of the development of porosity assumes that selective dissolution of the less noble metal creates a high vacancy concentration near the surface. The coalescence of these vacancies generates stresses, which may become sufficiently high to cause formation of cracks (c). The vacancy concentration near the surface increases with potential due to the higher dissolution rate, and so does the tensile stress. Eventually, at a sufficiently anodic potential this leads to cracking. The cracks expose new alloy surface and permit dissolution to continue. Figure 7.37 shows the cracked surface of a zinc-iron alloy (δ-phase, 6.9%Fe) after polarization in an acid sulfate solution under conditions where zinc selectively dissolved [15].

It is well possible, that several of the described mechanisms act together, their relative importance depending on the properties of the alloy and the prevailing corrosion conditions.

7.2.3 Intergranular corrosion

Because the electrochemical properties of grain boundaries often differ from those of the bulk metal, a corrosion cell can form between a grain boundary and the rest of the surface. The result is a local attack at the grain boundary known as ***intergranular corrosion***. In practice, intergranular corrosion is usually the result of inappropriate heat treatment, which leads to segregation or precipitation phenomena at the grain boundaries. Such heat treatments are referred to as ***sensitizing heat treatment***.

© 2007, first edition, EPFL Press

Intergranular corrosion is best known to occur with certain types of stainless steels and aluminum alloys, but in principle, many other alloys can be susceptible to this kind of corrosion damage.

Intergranular corrosion of austenitic stainless steel

The mechanism of intergranular corrosion of austenitic stainless steel has been extensively studied. Figure 7.38 shows the ***sensitization diagram*** for such a steel [16]. The ordinate gives the temperature of the heat treatment and the abscissa represents the logarithm of the treatment time. The curves indicate the limit between the sensitized and non-sensitized states. With their characteristic form, they resemble the TTT diagrams of carbon steel that describe the transformation of austenite into ferrite and cementite. The particular shape of the sensitization curves shown in the Figure 7.38 can be explained as follows. The sensitization of austenitic stainless steel is generally due to the *precipitation of chromium carbides* at the grain boundaries. The overall kinetics of the process depends on the rate of two phenomena. On one hand, the rate of nucleation increases with increasing supersaturation, thus with decreasing temperature. On the other hand, the rate of diffusion of carbon towards the grain boundaries increases with temperature. The time required to sensitize the alloy as a function of temperature therefore passes through a minimum.

Chromium carbide has the formula $Cr_{23}C_6$, but the carbides that precipitate at grain boundaries normally contain iron as well. Their formula is therefore usually written as $(Cr,Fe)_{23}C_6$. Because of their larger size, chromium atoms diffuse more slowly in steel than carbon atoms. As a consequence, the precipitation of chromium carbides at the grain boundary causes a depletion of chromium in the immediate vicinity (Figure 7.39). Indeed, under the specific thermal conditions of the sensitizing heat treatment, the diffusion of chromium is too slow to reestablish a uniform chromium concentration throughout. A low chromium concentration is known to render passivation more difficult (Chapter 6). The depleted zones near the grain

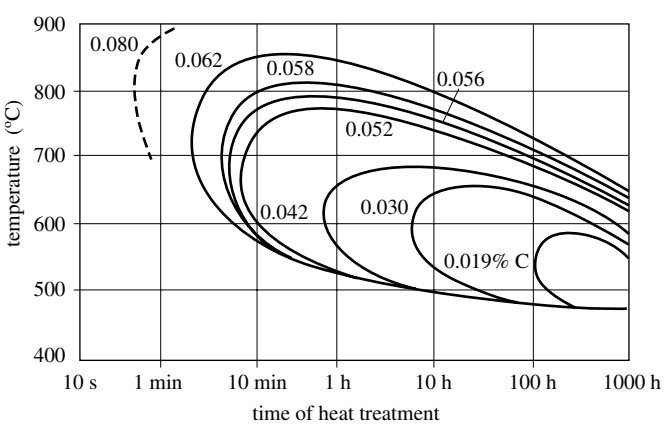


Figure 7.38 Sensitizing of 304 stainless steels of different carbon content [16].

© 2007, first edition, EPFL Press

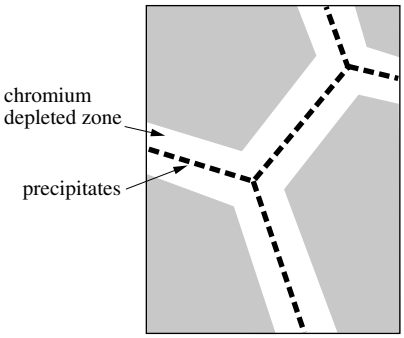


Figure 7.39 Chromium depletion at grain boundaries due to the precipitation of chromium carbides (schematic).

boundaries therefore passivate less easily than the rest of the surface. Under certain conditions, they may be active, while the grain cores remain passive. A corrosion cell then forms, with the depleted zone acting as the anode and the grain core as the cathode. The small anode/cathode surface ratio thus leads to rapid corrosion of the anodic zones.

The sensitivity of stainless steel to intergranular corrosion is not always due to chromium carbide precipitation though. Certain thermal treatments cause the precipitation phosphorous compounds such as $(Cr,Ni,Fe)_3P_2$ or of intermetallic phases at the grain boundaries. For example, in molybdenum-containing austenitic stainless steels, the intermetallic $(Fe,Cr,)Mo_2$ may precipitate at grain boundaries or the metallic σ phase, containing Fe, Cr, Ni and Mo. At high potential, near the transpassive region, when these or similar precipitates dissolve, they cause a selective attack at the grain boundary.

We conclude from this discussion that two mechanisms can lead to intergranular corrosion. They are schematically represented in Figure 7.40.

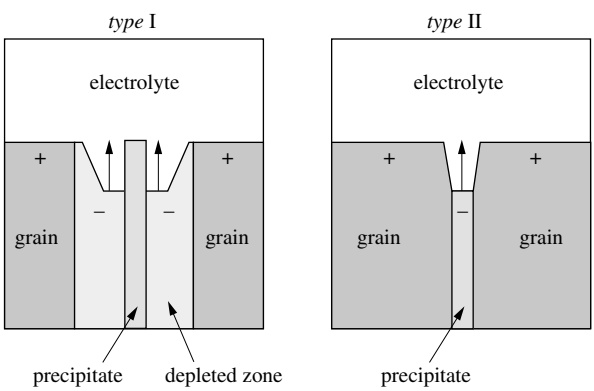


Figure 7.40 Mechanisms of intergranular corrosion. Type I: zone depleted of passivating metal. Type II: anodic precipitate.

© 2007, first edition, EPFL Press

In the first mechanism (*type I*), the precipitate is inert or even cathodic with respect to a zone immediately adjacent, which is depleted in elements that promote passivation. The depleted zone being anodic with respect to the rest of the surface it selectively corrodes. Type-I intergranular corrosion is particularly obnoxious under conditions where the zone adjacent to the grain boundary becomes active, while the remainder of the surface remains passive. Intergranular corrosion of austenitic stainless steel due to the precipitation of chromium carbide is the best-known example of type I intergranular corrosion.

In the second mechanism (*type II*), the precipitate acts as the anode and preferentially dissolves. Generally, dissolution of intermetallic or non-metallic precipitates requires relatively high potentials and is limited to precipitates that are less stable towards oxidizing conditions than the (normally passive) base metal.

The same metal can be sensitive to both types of intergranular corrosion depending on the prevailing potential. The susceptibility to intergranular corrosion of a given alloy such as austenitic stainless steel, therefore not only depends on the thermal treatment but also on the conditions to which it is exposed: potential, temperature, anions.

Intergranular corrosion tests

There are a number of tests available for the evaluation of the sensitivity of a given stainless steel to intergranular corrosion (Table 7.41). They differ in severity of corrosion conditions and therefore do not reveal the same phenomena. The ***oxalic acid test*** consists of an electrochemical attack at constant current. The morphology of the corroded surface is then compared with reference samples. The main advantage of this technique is its rapidity (the attack takes about one minute). The applied potential lies in the transpassive region, where the chromium easily dissolves. The oxalic acid test therefore does not reveal the zones depleted in chromium, but rather exposes the presence of carbides and intermetallic phases.

Table 7.41 Common testing methods for intergranular corrosion [17].

Method	Solution	Conditions	Evaluation method
Oxalic acid	Oxalic acid 10%	Application of a constant anodic current	Optical microscopy
Strauss test	H_2SO_4 +$CuSO_4$ +Cu	Immersion in boiling solution	Observation of cracks upon bending
Huey test	Concentrated HNO_3	Immersion in boiling solution	Mass loss
Streicher test	H_2SO_4 +$Fe_2(SO4)_3$	Immersion in boiling solution	Mass loss

The other testing methods mentioned in Table 7.41 are simple immersion tests carried out at elevated temperature, most often in boiling solution. Established norms

© 2007, first edition, EPFL Press

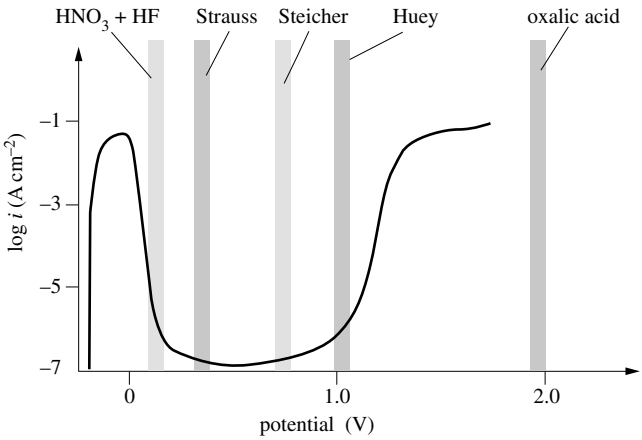


Figure 7.42 Corrosion potentials of different intergranular corrosion tests [18,19].

determine the exact conditions of the test. A major drawback of immersion testing is the time required. It can last from several hours to several days, depending on how aggressive the solution is.

The ***Strauss test*** employs a solution of H_2SO_4 + $CuSO_4$, to which, most often, copper filings are added. A corrosion potential close to the active-passive transition is established, which makes this test sensitive to chromium depletion at the grain boundaries. The ***Streicher test*** uses a solution of ferric sulfate, and the corrosion potential is therefore higher than in the previous testing method. The ***Huey test*** uses yet another solution: concentrated nitric acid. The corrosion potential is thus situated at the boundary of the passive and transpassive regions. This method leads to preferential attack of intermetallic precipitates at grain boundaries. Figure 7.42 gives the corrosion potentials corresponding to each of these different tests. For comparison, it also includes the polarization curve of an austenitic stainless steel measured in 1M sulfuric acid at 90 °C .

Intergranular corrosion near welds

Austenitic stainless steels become susceptible to intergranular corrosion after a sensitizing heat treatment at temperatures between 500 and 800 °C (Fig. 7.38). Sensitization can happen accidentally in the vicinity of welds or during the slow cooling of an object having been heat treated at temperatures in excess of 1000 °C well beyond the sensitizing range. Figure 7.43 shows schematically the extent of local heating that occurs near a weld and that may be viewed as the equivalent of a non-uniform heat treatment of the alloy. If the conditions of temperature and time for a given location correspond to the sensitization region, this zone becomes susceptible to intergranular corrosion. Such sensitized zones are normally found at a certain distance from the welds.

© 2007, first edition, EPFL Press

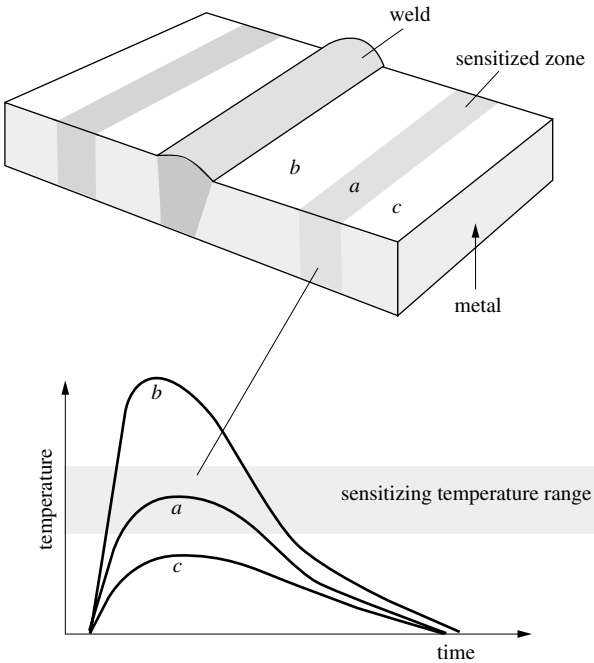


Figure 7.43 Intergranular corrosion near a weld and the variation of the temperature as a function of the distance and time (schematic).

In practice intergranular corrosion of austenitic stainless steels near welds can be avoided in a number of ways:

- by the use of stainless steels with low-carbon content (L type, < 0.03 % C);
- by a homogenizing heat treatment of the welded piece;
- by the use of a ***stabilized stainless steel*** that contains a small amount of carbide forming elements such as Ti or Nb.

The carbides of these elements are more stable than those of chromium and precipitate at temperatures where chromium remains in solution. By a heat treatment above the sensitizing region one can thus precipitate the dissolved carbon and prevent the formation of chromium carbide at the grain boundaries.

Intergranular corrosion of ferritic stainless steels

Ferritic stainless steel has the reputation of being less sensitive to intergranular corrosion than austenitic stainless steel. This type of corrosion can nevertheless take place under certain conditions of thermal treatment [20]. The diffusion coefficients of both carbon and chromium in ferrite are larger than in austenite. Grain boundary precipitation of carbides and nitrides of chromium can therefore occur at temperatures of 540–600 °C already. The behavior differs from that of austenitic stainless steel, which becomes sensitized at higher temperatures only. Because of the larger diffusion

© 2007, first edition, EPFL Press

coefficient, a heat treatment at temperatures near 800 °C permits to homogenize the chromium concentration, thus preventing intergranular corrosion.

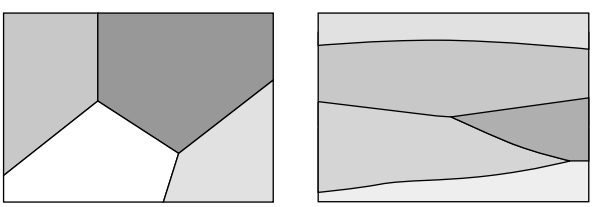
Intergranular corrosion of aluminum alloys

Certain aluminum alloys are susceptible to intergranular corrosion [21]. Al-Cu alloys, for example, harden by precipitation of an Al_2Cu phase (θ phase). Under certain conditions, in particular during annealing near 200 °C, precursors of the θ phase selectively precipitate at the grain boundaries, where nucleation is energetically more favorable. The precipitates being more noble than the matrix, they may form a corrosion cell with the copper depleted region in their immediate vicinity acting as anode. When exposed to a corrosive environment, the alloy may thus become subject to type-I intergranular corrosion.

Table 7.44 Galvanic effects in precipitation hardened aluminum alloys.

Designation of alloy family	Composition	Precipitate	Precipitate potential with respect to alloy
2000	Al-Cu Al-Cu-Mg	Al_2Cu Al_2CuMg	more noble
5000	Al-Mg	Mg_5Al_8	less noble
6000	Al-Mg-Si	Mg_2Si, Si	slightly more noble
7000	Al-Zn-Mg Al-Zn-Mg-Cu	$MgZn_2$ $Mg_3Zn_3Al_2$	less noble

Other aluminum alloys are susceptible to type-II intergranular corrosion. Under certain conditions of thermal processing (200–250 °C), $MgZn_2$ intermetallics precipitate at the grain boundaries of Al-Zn-Mg alloys. The corrosion potential of these precipitates is less noble than that of the matrix. When corrosion cells form between the precipitates and the matrix, the precipitates act as the anode. The resulting attack is particularly damaging when the precipitate is in the active state, while the alloy remains passive. Such situations can be found in presence of chloride when the precipitate has a lower pitting potential than the matrix [22]. Above its pitting potential the precipitate actively dissolves leading to intergranular corrosion of the alloy. Table 7.44 gives an overview of the electrochemical properties of the most common precipitates found in commercial aluminum alloys.



(a) (b)

Figure 7.45 Grain structure of (a) recristallized metal with equiaxed grains; and of (b) cold rolled metal with elongated grains.

© 2007, first edition, EPFL Press

Exfoliation corrosion of aluminum alloys is a special case of intergranular corrosion that is observed on alloys with elongated grain structure resulting from cold rolling (Figure 7.45). It manifests itself as a grain boundary attack that follows the orientation of the texture. As a consequence, flat grains may separate like foils. Exfoliation corrosion is related to the precipitation of intermetallic phases at grain and subgrain boundaries. Its mechanism may involve type-I or type-II intergranular corrosion.

7.3 PITTING CORROSION

7.3.1 Phenomenological Aspects

Pitting corrosion refers to a localized corrosive attack of a passive metal that leads to the development of small cavities (pits) (cf. Figure 6.38). Their growth eventually leads to perforation of a metal plate or pipe wall. Pitting corrosion damage, therefore, can be quite important, even if the absolute amount of corroded material is small. The number and shape of corrosion pits varies widely. Three typical pit shapes are presented in Figure 7.46: (a) deep narrow pits exhibiting crystalline attack; (b) occluded pits growing under a cover constituted of corrosion products or non-corroded metal; (c) hemispherical pits with a polished surface. This last morphology often appears during potentiostatic tests in concentrated electrolytes. The size of corrosion pits varies typically from a few micrometer to some millimeters.

Pitting corrosion requires the presence of ***aggressive anions***, most often chloride ions, and of an oxidizing agent such as oxygen or ferric ions. A corrosion cell forms between the growing pit which is the anode and the passive surface surrounding the pit which serves as the cathode (Figure 7.47). Because the anode/cathode surface ratio is small, dissolution inside the pit can be quite rapid.

For characterizing the corrosion damage due to pitting one measures the depth distribution of pits. From such data the evolution of maximum pit depth with time or the time to perforation can be estimated using statistical methods [23,24]. Because pitting is a random process the observed pit depth depends on the surface area taken into account. The depth L has been found to vary according to a power law (7.25) or a logarithmic law (7.26) where A is the surface area and K_i (i = 1,2 . .) are constants.

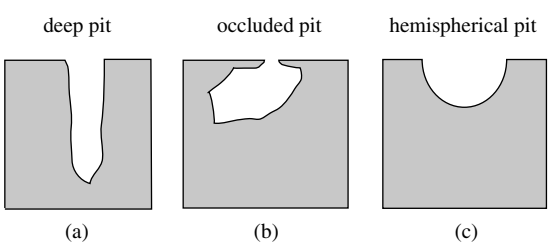


Figure 7.46 Shape of corrosion pits: (a) deep pit; (b) occluded pit; and (c) hemispherical pit.

© 2007, first edition, EPFL Press

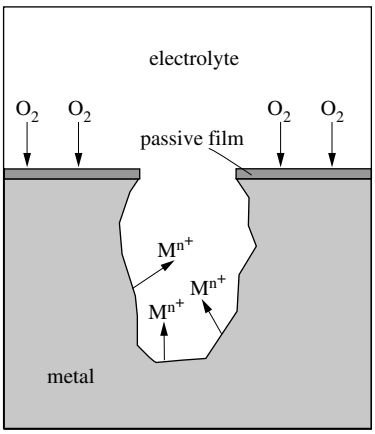


Figure 7.47 Partial reactions in pitting corrosion.

$$L(A) = K_1 \cdot A^{K_2} \tag{7.25}$$

$$L(A) = K_3 + K_4 \log A \tag{7.26}$$

These relationships can be interpreted in terms of extreme value statistics by assuming different types of distribution laws [23].

Critical pitting potential

Pitting corrosion takes place when the corrosion potential exceeds the ***critical pitting potential*** (or simply *pitting potential*) E_b, which limits the region of stable passivity at the anodic end (Chap. 6). Figure 7.48 illustrates situations of stable passivity (a) and of pitting (b) by means of Evans diagrams: if $E_{cor} < E_b$ passivity is stable and if $E_{cor} > E_b$ pitting corrosion occurs.

For thermodynamic reasons, the corrosion potential is always lower than the reversible potential of the oxidizing agent ($E_{cor} < E_{rev,oxidant}$). Dissolved oxygen is the

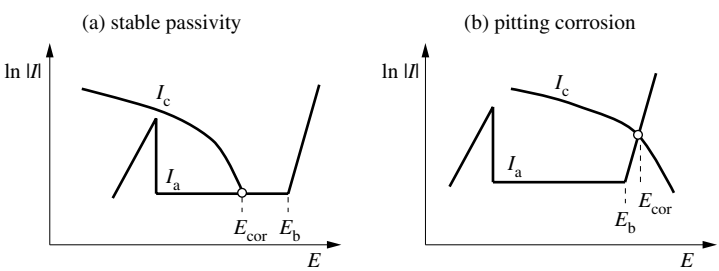


Figure 7.48 Evans diagram of a passive electrode showing the sensitivity to pitting corrosion (schematic): (a) $E_{cor} < E_b$; and (b) $E_{cor} > E_b$.

© 2007, first edition, EPFL Press

most common oxidizing agent present in natural environments. Under these conditions the reversible potential of oxygen, which varies with pH, limits the maximum value of the corrosion potential ($E_{cor} < E_{rev,O2}$). According to Figure 7.48, if the pitting potential is higher than this value, pitting corrosion cannot take place. In natural environments, a given metal is therefore resistant to pitting corrosion if condition (7.27) is fulfilled:

$$E_b > 1.23 - 0.059\, pH \quad \text{(V)} \tag{7.27}$$

Pitting corrosion tests

There are two ways to study the resistance of a metal to pitting corrosion:

- immersion testing in the presence of an oxidizing agent;
- electrochemical determination the pitting potential.

Most metals resist corrosion less well as the temperature is increased. Immersion testing for pitting corrosion makes use of this observation. The sample is immersed into a solution of $FeCl_3$, whose temperature is gradually increased. The highest temperature at which the metal resists pitting corrosion (observed visually or by chemical analysis) is called the ***critical pitting temperature***. It can be used for ranking different metals in terms of their resistance to pitting corrosion. One should keep in mind, however, that this test does not allow for a comparison of the influence of different environments and it gives no mechanistic information.

In electrochemical experiments for the determination of the pitting potential one either controls the potential (potentiodynamic method) or, more rarely, the current (galvanostatic method). The composition and temperature of the electrolyte are selected such as to represent the real environment to which the metal will be exposed, but without the oxidant present, whose effect is simulated by the anodic polarization.

Figure 7.49(a) describes the principle of the ***potentiodynamic method*** for the determination of the critical pitting potential. At the pitting potential the anodic current increases sharply, providing a simple measure of E_b. Unfortunately, the value thus obtained depends on sweep rate, it is generally the higher the faster one sweeps the potential in the anodic direction. To get representative values of E_b one therefore must work at slow sweep rates corresponding to the measurement of pseudo-steady-state polarization curves (Chapter 5). To avoid these problems some authors have proposed to use the ***repassivation potential*** rather than the pitting potential for characterizing the resistance to pitting corrosion. When the sweep direction is reversed, one observes a hysteresis in the current; the potential at which the metal becomes passive again (repassivation potential) is lower than the potential E_b corresponding to passivity breakdown during the forward scan. However, the size of the hysteresis depends on the experimental parameters, notably the sweep rate and the potential of sweep reversal which determine the extent of dissolution before the metal repassivates. The usefulness of reverse sweep experiments for pitting corrosion testing is therefore questionable.

© 2007, first edition, EPFL Press

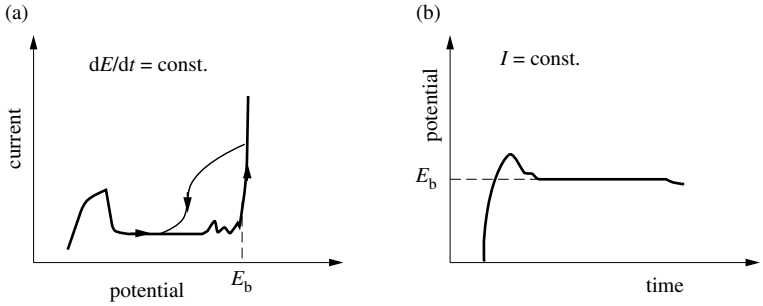


Figure 7.49 Experimental determination of pitting potential: (a) by the potentiodynamic method and (b) by the gavanostatic method.

The ***galvanostatic method*** consists of the application of a constant current that exceeds the passivation current and recording of the resulting potential transient (Figure 7.49(b)). Upon current switch-on, the potential increases as long as the passive film remains intact. After a certain time, which depends on the applied current density, film breakdown leads to dissolution by pitting. As a consequence, the anodic potential decreases and finally stabilizes when the rate of pit growth becomes steady.

7.3.2 Factors that influence the pitting potential

The critical pitting potential is not a well-defined kinetic or thermodynamic quantity. Its value depends as much on the rate of pit initiation as on the kinetics of pit growth and repassivation. In addition, the measurement of the pitting potential of industrial alloys are often not well reproducible, requiring statistical analysis to determine the probability of pitting for a given potential. Despite these restrictions, the pitting potential is a useful quantity for the characterization of the susceptibility of metals and alloys to pitting corrosion. The pitting potential as a property of the metal-environment system, depends on a number of factors:

- the chemical composition and microstructure of the metal;
- the surface state and the presence of inclusions;
- the chemical composition of the electrolyte, in particular the concentrations of the aggressive and non-aggressive anions;
- the temperature;
- the prevailing convection conditions.

Influence of alloy composition

Table 7.50 shows the pitting potentials of different metals and alloys observed in 0.1 M NaCl at 25 °C. The particularly high values for titanium and tantalum provide an explanation for their excellent resistance to pitting corrosion in chloride-containing environments.

© 2007, first edition, EPFL Press

Table 7.50 Pitting potentials of different metals in 0.1 M NaCl, 25 °C [25]

Metal	E_b (V)
Al	–0.4
Fe-12 Cr	0.2
Fe18Cr-8Ni	0.3
Ni	0.3
Fe-30Cr	0.6
Ti	~ 10
Ta	~ 25

The pitting potential of alloys depends on their composition. For stainless steels, its value increases with increasing chromium content (Figure 7.51) [26]. At sufficiently high chromium concentration, dependent of the environment, pitting no longer occurs and uniform transpassive dissolution is observed. Under the test conditions shown in Figure 7.51, this change occurs at about 40% chromium.

To improve their resistance to pitting corrosion, molybdenum is often added to stainless steel. Figure 7.52 illustrates the effect of molybdenum on the pitting potential of two stainless steels in NaCl [27]. Certain other alloying elements, in particular nitrogen and tungsten, also are beneficial; whereas others, such as copper and sulfur, reduce the resistance to pitting corrosion.

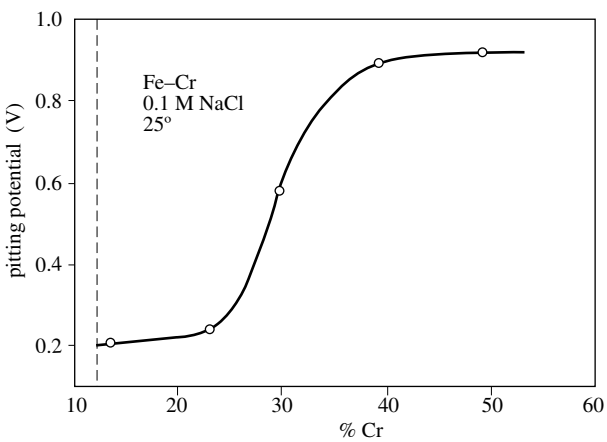


Figure 7.51 Pitting potential of Fe-Cr alloys as a function of Cr content [26].

Influence of inclusions

Pitting of industrial metals and alloy almost always starts at non metallic inclusions, notably sulfides, or intermetallic precipitates. It is therefore not surprising that the presence of inclusions or precipitates on a metal surface lowers the value of the pitting potential. As an illustration, Figure 7.53 presents anodic polarization curves

© 2007, first edition, EPFL Press

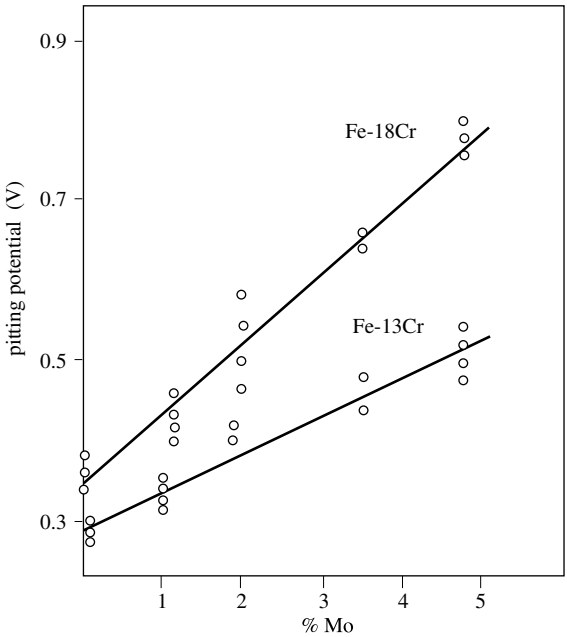


Figure 7.52 Effect of molybdenum on pitting potential of two Fe-Cr alloys in 1 M NaCl, 25 °C [27].

in HCl of two ferritic stainless steels, one of industrial grade and the other of high purity. The high purity alloy has a significantly wider passive region [28].

Inclusions or intermetallic precipitates can facilitate pit initiation and growth in different ways. Figure 7.54 shows the effect of inclusions that are anodic, cathodic or inert with respect to the base metal. Inert inclusions (a) do not electrochemically interact with the base alloy, but they can still play a role in the initiation of pitting: if

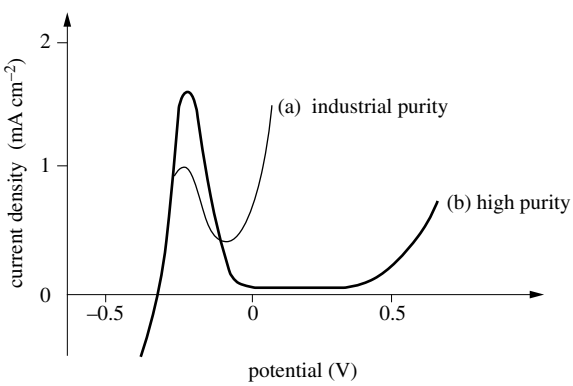


Figure 7.53 Anodic polarization curves of both (a) industrial grade and (b) high purity samples of Fe-17Cr in 0.12 M HCl at 25 °C [28].

© 2007, first edition, EPFL Press

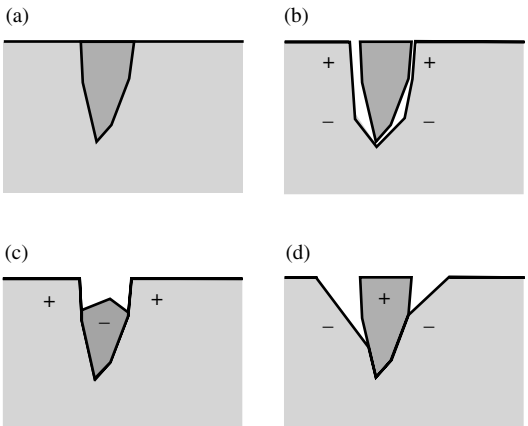


Figure 7.54 Effect of inclusions on pit initiation: (a) unattacted surface with an inclusion; (b) formation of an aeration cell in the space between an inert inclusion and the metal; (c) anodic dissolution of the inclusion; (d) formation of a corrosion cell between the cathodic inclusion and the metal.

the thermal expansion coefficient of the inclusion differs from that of the base metal, cracks between the inclusion and the matrix can develop during quenching. A microscopic aeration cell (b) can then develop and lead to formation of a pit by undercutting via a crevice corrosion mechanism. Inclusions that are less noble than the base metal (c) become the anode of a corrosion cell and dissolve, creating a pit nucleus. Oppositely, cathodic inclusions (d) favor the dissolution of the base metal in their immediate vicinity. They eventually fall off, leaving a small pit on the surface.

Influence of electrolyte composition

Generally speaking, a higher concentration of aggressive anions goes hand in hand with a lower value of the pitting potential. Figure 7.55 shows the variation of the pitting potential of aluminum as a function of the NaCl concentration in the electrolyte [29]. One observes a linear decrease of E_b with the logarithm of concentration. Similar relationships have been observed with other metals and alloys:

$$E_b = A' - B' \log c_{Cl^-} \tag{7.28}$$

where A' and B' are constants. At ambient temperature, B' is about 0.1V. For comparison, pitting potentials measured in 1M solutions of NaBr and NaI are also given in Figure 7.55. Their value increases in the order Cl < Br < I. However, this trend is not universal: for titanium, for example, the pitting potential in bromide solution at ambient temperature is about 2 V, while in chloride solution it is typically 10 to 12 V.

© 2007, first edition, EPFL Press

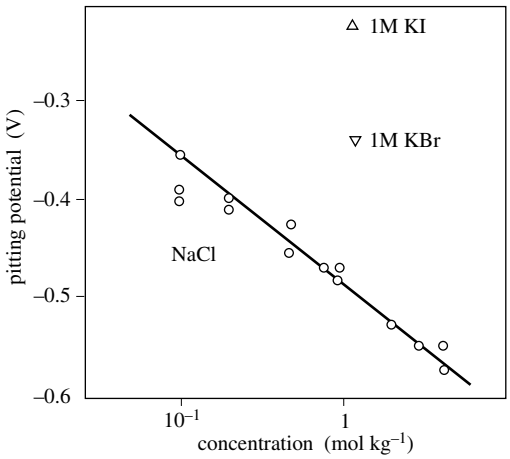


Figure 7.55 Effect of Cl^- concentration on pitting potential of aluminum at 25 °C. Also indicated is the pitting potential in 1M KBr and in 1M KI [29].

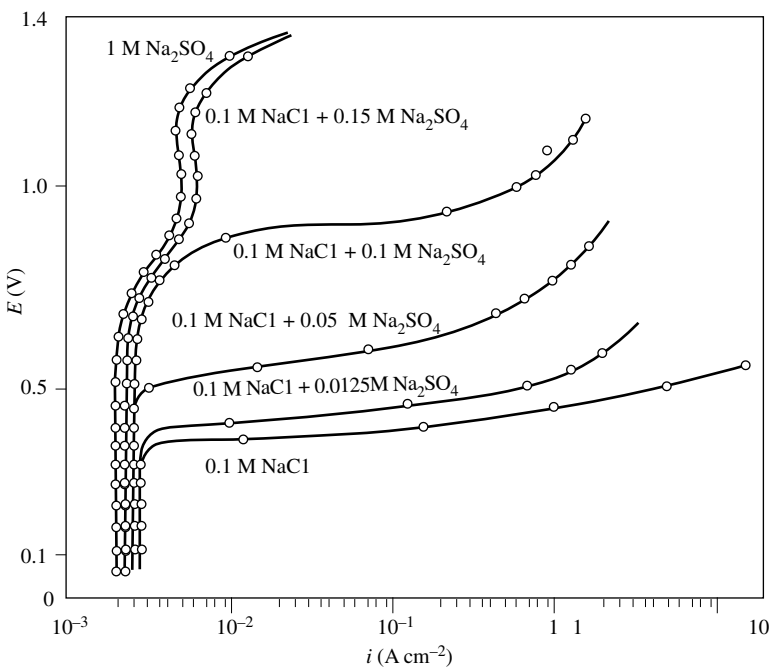


Figure 7.56 Potentiostatic polarization curves showing effect of electrolyte composition on pitting behavior of Fe-18Cr-8Ni exposed to Cl^- and SO_4^{2-} anions [25].

© 2007, first edition, EPFL Press

In the presence of both aggressive (Cl^-, Br^-, etc) and non-aggressive (OH^-, NO_3^-, SO_4^{2-}, etc.) anions, the pitting potential depends on the ratio between the two. For chloride containing electrolytes one thus finds:

$$E_b = A' - B' \log \frac{c_{Cl^-}}{\sum c_i} \tag{7.29}$$

where Σc_i represents the sum of the non-aggressive anion concentrations. Figure 7.56 shows potentiostatic polarization curves, measured on an 18Cr-8Ni stainless steel, in different chloride-sulfate mixtures. The results clearly show that the pitting potential depends on the concentrations of both kinds of ions.

Influence of temperature

Figure 7.57 shows the variation of the pitting potential as a function of the temperature for three different stainless steels, each immersed into 3% NaCl. For these three alloys, as the temperature increases, E_b decreases. This behavior explains the origin of a critical pitting temperature in potential-controlled immersion tests in the presence of an oxidizing agent. The pitting temperature corresponds to that temperature at which the potential E_b becomes equal to or less than the corrosion potential. Between the three alloys the value of E_b increases in the order Fe-17-Cr < Fe-18Cr-10Ni < Fe-17Cr-13Ni-2Mo, confirming the beneficial effect of alloy elements on pitting.

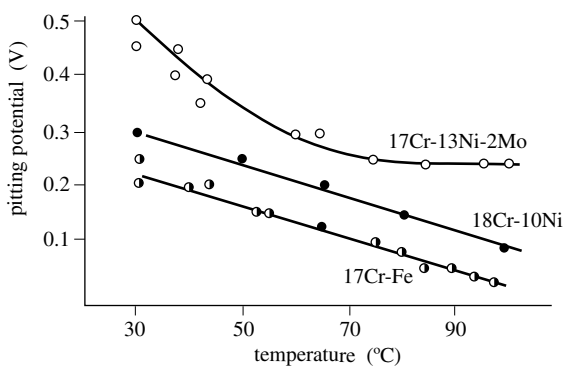


Figure 7.57 Effect of temperature on pitting potential of stainless steels in 3% NaCl [30].

7.3.3 Metastable pits

The study of pitting corrosion of engineering metals and alloys involves three aspects [31]:

- pit nucleation;
- metastable pitting;
- stable pit growth.

© 2007, first edition, EPFL Press

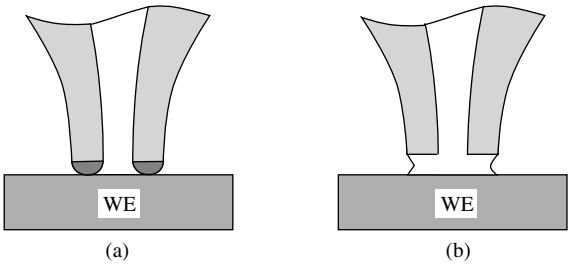


Figure 7.58 Microelectrochemical methods for the study of localized corrosion: (a) capillary with rubber seal, (b) capillary with hanging drop of electrolyte.

In Section 6.3.2 different mechanisms of passive film breakdown and pit nucleation were discussed. There is strong evidence that on engineering materials pits normally develop at non-metallic inclusions or intermetallic precipitates, depending on the alloy. Since the number and distribution of these defect sites are difficult to control, pit initiation is usually not a well reproducible phenomenon. In a potentiodynamic experiment, as one increases the anodic potential, an ever greater number of inclusions and precipitates may become electrochemically active and start to dissolve, thus permitting nucleation of a pit. A pit nucleus may then react in different ways: it may immediately passivate (repassivation); it may grow for a limited time before repassivating (metastable pit formation) or it can develop into a pit that keeps growing with time (stable pit growth).

To study these phenomena ***microelectrochemical methods*** have been found particularly useful [32,33]. The principle is shown in Figure 7.58. A glass capillary filled with electrolyte and containing a rubber seal is tightly placed on a metal surface thus limiting the surface of the working electrode. The back of the capillary is connected to the reference and counter electrodes. Another possibility is to use a hanging drop instead of a rubber seal to limit the anode area. The described techniques permits to work with electrodes with a diameter as small as 50 micrometer. One can thus study the effect of inclusions on the local anodic behavior of a metal.

As an illustration, Figure 7.59 shows potentiodynamic polarization curves measured on an industrial stainless steel containing 0.003% sulfur in a sulfate solution using, respectively, a conventional 10 mm diameter electrode (a) and a 150 micrometer diameter microcapillary (b). On the microelectrode small current spikes are seen in the passive potential region. They result from electrochemical reactions of sulfide inclusions. On the large size electrode these spikes are hidden in the background current and therefore not visible. In the sulfate solution used, the stainless steel does not pit. In spite this one observes current spikes in the passive potential region. The results indicate that pits nucleate, but immediately repassivate. Figure 7.60 compares the behavior in sulfate and in chloride solution. Similar current spikes occur in both solutions, but they are much more pronounced in presence of chloride, indicating that pit nuclei grow to a larger size before they repassivate. Furthermore, the alloy exhibits a critical pitting potential above which pits continue to grow. The results suggest that

© 2007, first edition, EPFL Press

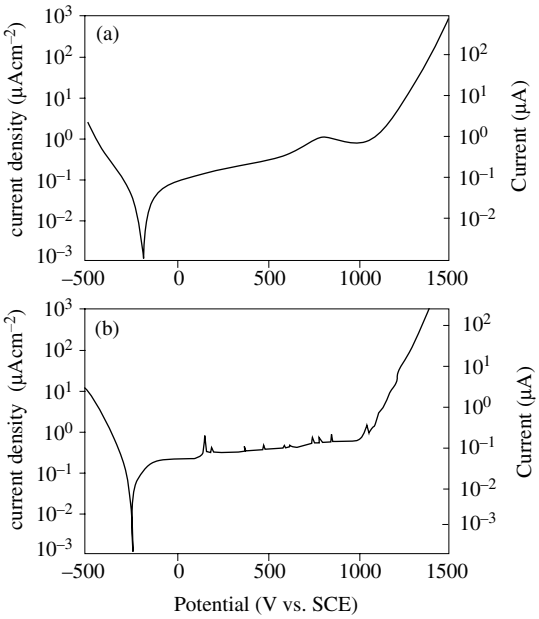


Figure 7.59 Potentiodynamic polarization curves of 304 stainless steel (0.003% S) in 1M Na_2SO_4 measured at a scan rate of 0.2 mV/s on (a) an electrode of 10 mm diameter and (b) on a microelectrode of 150 μm diameter [34].

chloride ions favor pitting by rendering repassivation of pit nuclei more difficult. In addition, they could possibly affect the dissolution behavior of the inclusions.

Metastable pitting may be considered as a precursor of stable pit growth which occurs above the pitting potential. The charge associated with the formation and repassivation of metastable pits can be calculated by integrating the corresponding current spikes. This allows one to estimate the extent of pit growth before they

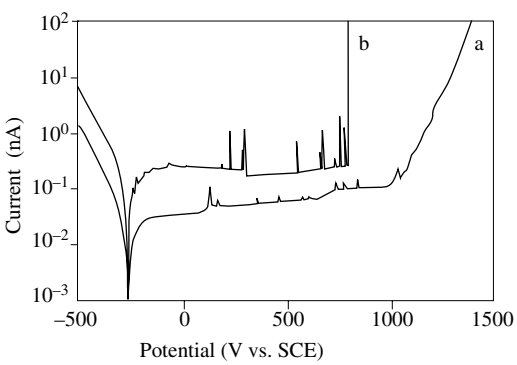


Figure 7.60 Potentiodynamic polarization curves of 304 stainless steel (0.003% S) measured on microelectrodes at a scan rate of 0.2 mV/s: (a) 1M Na_2SO_4, electrode diameter 150 μm, (b) 1M NaCl, electrode diameter 220 μm [34].

© 2007, first edition, EPFL Press

repassivate. In principle one can differentiate a pit nucleus from a metastable pit by the amount of charge passed before repassivation. However, there is no clear limit between the two, and the two terms are often used interchangeably.

Whether a pit can grow or repassivates depends not only on the prevailing potential, but also on the local chemistry in the pit nucleus, which differs from that of the bulk electrolyte, because the dissolution reaction causes concentration changes such as anion accumulation and a lowering of the pH. Furthermore, the chemical nature of the inclusions as well as their size and shape affect repassivation. For example, sulfide inclusions are well known pit initiation sites. Sulfides can anodically dissolve at potentials well below the pitting potential of many alloys and their dissolution products can inhibit passivation by adsorbing at the metal surface (Chapter 6). For this reason sulfide inclusions are more damaging than electrochemically inert inclusions such as nitrides.

The described pit-initiation mechanism involving inclusions explains the fact that measured pitting potentials for a given alloy-environment combination usually exhibit wide scattering and that the values obtained depend on the surface area considered. Indeed, when inclusions of different size are randomly distributed on a surface, pit initiation is expected to occur randomly also. To evaluate the probability of pitting corrosion from pitting potential measurements in the laboratory, statistical methods are therefore needed. Figure 7.61 shows a probability plot of pitting potentials of different ferritic stainless steel measured in potentiodynamic experiments performed under identical conditions except for the alloy composition. While all alloys show a scattering of pitting potentials, the pitting probability at a given potential clearly varies with the alloy composition.

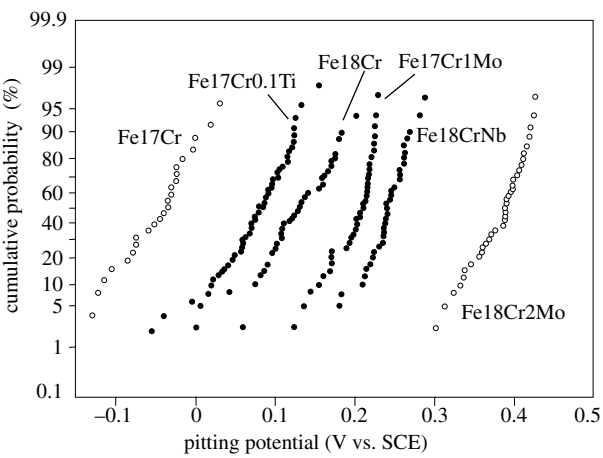


Figure 7.61 Normal probability plot of the pitting potentials of different ferritic stainless steels measured by potentiodynamic polarization in 3.5% NaCl at 30°C under nitrogen atmosphere. Potential scan rate: 16.7 mV min^{-1} [23].

© 2007, first edition, EPFL Press

7.3.4 Pit growth rate

The rate of pit growth is proportional to the local rate of anodic dissolution. Several phenomena may be rate limiting: the potential difference between anode and cathode, the charge transfer kinetics of the dissolution reaction, the rate of cathodic reduction of the oxidant, the mass transport in the pit. Taking into consideration all these aspects in a realistic pit growth model is extremely complicated. We shall limit the discussion to two limiting cases.

- growth controlled by the ohmic resistance in the electrolyte;
- growth controlled by the transport of the dissolution products.

To highlight the two mechanisms, we use the one-dimensional model presented in Figure 7.62. We assume that the walls are insulating and that uniform dissolution takes place at the bottom of the growing pit.

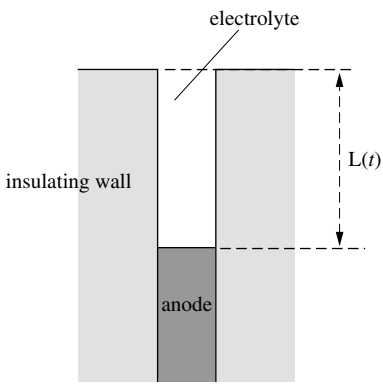


Figure 7.62 One dimensional model of a corrosion pit.

Growth under ohmic control

When the anodic current is under ohmic control and there are no concentration gradients, the following relation describes the dissolution rate at the base of the pit:

$$i_a = \kappa \frac{\Delta\Phi}{L(t)} \tag{7.30}$$

Here, κ represents the electrolyte conductivity in the corrosion pit, $L(t)$ is the depth of the pit that changes with time, and $\Delta\Phi$ corresponds to the potential difference between the pit base and the pit opening. For relatively deep corrosion pits, the potential drop in the electrolyte outside the pit is negligible compared to that within the pit because of the much larger cross section for current flow. In this case, $\Delta\Phi$ is approximately equal to the potential difference between anode and cathode. According to Faraday's law, the growth rate of the pit is proportional to the anodic current density:

$$\frac{\mathrm{d}L(t)}{\mathrm{d}t} = \frac{i_a M}{n F \rho} \tag{7.31}$$

© 2007, first edition, EPFL Press

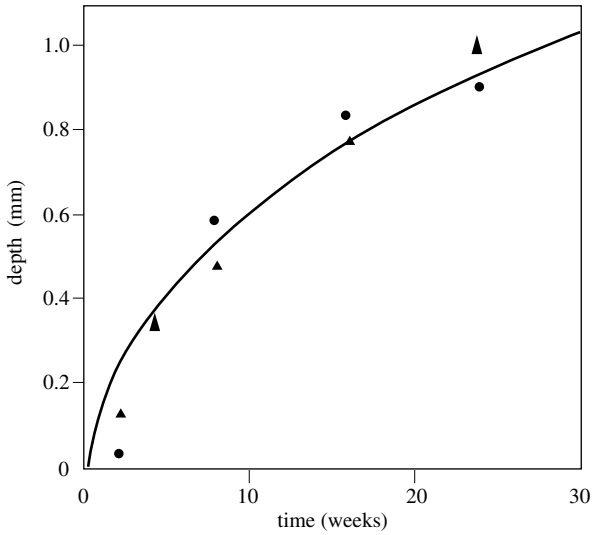


Figure 7.63 Pit depth on aluminum versus exposure time in natural water. The theoretical curve is calculated with equation (7.35) using $\kappa = 5.7 \times 10^{-4}\ \Omega^{-1}\ \text{cm}^{-1}$, $\Delta\Phi = 15$ mV. The different symbols represent different measurements carried out under identical conditions [35].

Here, M represents the atomic mass of the metal and ρ is its density. With (7.30) and (7.31):

$$L(t) = K_\Omega\, t^{1/2} \tag{7.32}$$

$$K_\Omega = \left(\frac{2\, M\, \kappa \Delta\Phi}{n\, F \rho} \right)^{1/2}$$

The depth of a one-dimensional pit increases with the square root of time. The proportionality constant K_Ω depends on the imposed potential difference and the electrolyte conductivity. Pit growth under ohmic control is observed mostly in weakly conducting electrolytes, such as drinking water. Figure 7.63 shows the progression of pit depth, measured on aluminum in contact with natural water [35]. The results are in good agreement with the simple theory developed above.

Mass-transport controlled growth

The metal ions produced at the dissolving surface in the pit are transported towards the exterior by diffusion and migration. Inside a corrosion pit, convection is generally negligible. A concentration gradient of metal ions builds up while metal ions diffuse outwards. To maintain electric neutrality of the solution anions migrate in the opposite direction and accumulate in the pit. When the metal ion concentration at the anode surface reaches saturation a salt film precipitates (Chapter 4). The

© 2007, first edition, EPFL Press

dissolution rate under these conditions becomes entirely limited by the rate of transport of dissolving metal ions from the pit bottom to the exterior. Assuming that at the time $t = 0$, the dissolving metal ion concentration of the solution is zero ($c_b = 0$), the following relation describes the anodic current density, where c_{sat} is the saturation concentration of the metal ions and D_{eff} an effective diffusion coefficient:

$$i_a = n\,F\,D_{eff}\,\frac{c_{sat}}{L(t)} \tag{7.33}$$

With (7.31) and (7.33):

$$L(t) = K_D\ t^{1/2} \tag{7.34}$$

$$K_D = \left(\frac{2\,M\,D_{eff}\,c_{sat}}{\rho}\right)^{1/2}$$

Just as in the limiting case of ohmic control, the depth of the corrosion pit increases with the square root of the time. However, the proportionality constant here differs, it includes the saturation concentration rather than the potential and the diffusion coefficient rather than the electrolyte conductivity. Figure 7.64 shows experimental results for potentiostatic anodic dissolution of nickel in chloride solution. The geometry of the electrochemical cell corresponds to the one-dimensional pit model represented in Figure 7.62. The results show that, after a certain time, the current density decreases according to the square root of the polarization time, independent of potential. Mass-transport controlled growth of corrosion pits is favored by a highly electrolyte conductivity and weakly soluble dissolution products.

An interesting application of mass transport controlled pit growth is found in electrochemical micromachining. When anodically dissolving a metal such as titanium

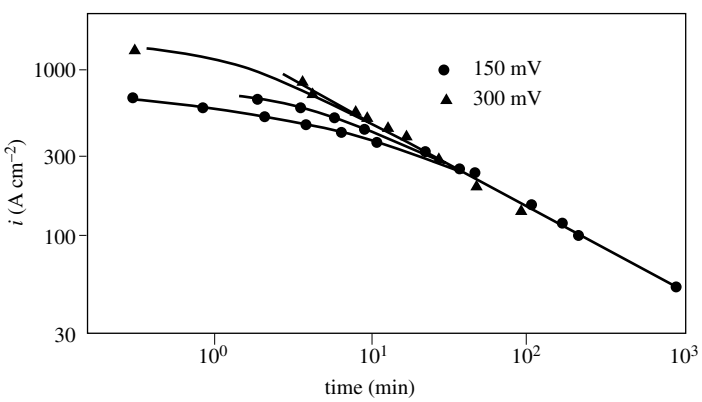


Figure 7.64 Potentiostatic dissolution of Ni in 0.1 M NaCl in a diffusion cell: variation of current density with time. The applied potentials are 150 mV amd 300 mV relative to the calomel electrode [36].

© 2007, first edition, EPFL Press

through a photoresist mask under conditions where dissolution is controlled by mass transport, isotropic dissolution leads to the development of well defined hemispherical cavities having a polished surface [37].

The growth behavior of real corrosion pits is more complicated than discussed here using a one-dimensional model based on ohmic or mass transport control. In a more general way, we can write for the variation of pit depth with time:

$$L(t) = K_p\, t^a \tag{7.35}$$

where K_p and a ($0 < a > 1$) are experimental constants having a value that depends on pit geometry and growth kinetics. The simple models discussed above provide a physical interpretation of K_p and a for two limiting cases.

Autocatalytic mechanism of pit growth

During pit growth, chloride ions accumulate in the pit by migration in order to preserve electric neutrality in the face of dissolving metal ions. Furthermore, dissolving metal ions undergo hydrolysis reactions, causing a lowering of pH inside the corrosion pit:

$$M^{n+} + H_2O = MOH^{(n-1)+} + H^+ \tag{7.36}$$

Both, a higher chloride concentration and a locally lower pH render repassivation of a growing pit more difficult. In addition, certain dissolution products from sulfide inclusions also inhibit repassivation. As the pit grows, the chemical composition of the electrolyte inside becomes thus increasingly aggressive. Pitting corrosion, therefore, is an *autocatalytic* process: once formed, the pit itself creates conditions that favor its further growth.

The autocatalytic nature of pit growth is responsible for the fact that the repassivation potential observed potentiodynamic experiments is generally lower than the breakdown potential (Figure 7.49(a)) and that it depends on the details of the experimental procedure: the concentration changes induced by the dissolution reaction during sweep reversal determine the ionic concentrations in the developing pit and the ease of repassivation.

Another observation can be explained by the autocatalytic nature of pit growth: renewing the solution inside a pit favors its repassivation. Because of this, open pits repassivate more easily that occluded pits. For the same reason, engineering structures are most exposed to pitting corrosion in situations involving a stagnant electrolyte when concentration gradients can build up undisturbed by convection, for example during prolonged shutdown of an installation.

General Bibliography

M. Fontana, *Corrosion Engineering*, 3rd edition, Mc Graw Hill International Edition, New York (1987).

© 2007, first edition, EPFL Press

H. H. Uhlig, R. W. Revie, *Corrosion and Corrosion Control,* 3rd edition, John Wiley, New York (1985).
H. Kaesche, *Corrosion of Metals,* Springer Verlag, Berlin (2003).
H. Isaacs, U. Bertocci, J. Kruger, S. Smialovska, editors, *Advances in Localized Corrosion*, NACE 9, Natl. Ass. Corrosion Engineers, Houston, Texas (1990).

References

[1] H. P. Hack, D. Taylor, *Metals Handbook*, 9th Edition, Vol. 13, ASM Metals, Park, Ohio (1987) p. 235.
[2] E. Budesvski, *Treatise of Electrochemistry,* Vol. 7, Plenum Press, New York (1983) p. 409.
[3] M. Datta, D. Landolt, *J. Electrochem. Soc.* 129, 1889 (1982).
[4] K. E. Heusler, D. Huerta, *J. Electrochem. Soc*. 136, 65 (1989).
[5] X. F. Yang, J. E. Castle, *Surf. Interface Anal.* 33, 894 (2002).
[6] C. Hubschmid, H. J. Mathieu, D. Landolt, *Surf. Interface Anal.* 20, 755 (1993).
[7] J. Erlebacher, M. J. Aziz, A. Kamura, N. Dimitrov, K. Sieradzki, *Nature* 410, 450 (2001).
[8] J. J. Laurent, D. Landolt, *Electrochim. Acta* 36, 49 (1991).
[9] D. Landolt, *Surface and Interface Anal.* 15, 395 (1990).
[10] H. Pickering, C. Wagner, *J. Electrochem Soc.* 114, 698 (1967).
[11] J. Erlebacher, *J. Electrochem. Soc.* 151, C614 (2004).
[12] S. Rambert, D.Landolt, *Electrochim. Acta 31*, 1421, 1441 (1986).
[13] H. Kaiser, G.A. Eckstein, *Encyclopedia of Electrochemistry*, Vol. 4, M. Stratmann, G. Frankel, editors, Wiley-VCH, Weinheim (2003) p.156.
[14] H. Kaesche, *Die Korrosion der Metalle*, 3rd edition, Springer Verlag (1990) p. 172.
[15] H. Kaiser, *Materials and Corrosion* 47, 34 (1996).
[16] *Metals Handbook*, 9th edition, Vol. 13, ASM Metals, Park, Ohio (1987) p. 551.
[17] M. A. Streicher, *Intergranular Corrosion of Stainless Alloys*, ASTM technical publication 656, R.F. Steigerwald, editor, Am. Soc. for Testing of Materials, Philadelphia (1978) p. 3.
[18] M. Henthorne, *Localized Corrosion*, ASTM technical publication 518, Am. Soc. for Testing of Materials, Philadelphia (1972) p. 108.
[19] G. Herbsleb, P. Schwab, *Werkstoffe und Korrosion* 37, 24 (1986).
[20] G. Béranger, J. C. Charbonnier, *Les aciers inoxydables*, P. Lacombe, B. Baroux, G. Béranger, editors, Les éditions de physique, Paris (1990) pp. 411-434.
[21] N. Birblis, R. G. Buchheit, *J. Electrochem. Soc.* 152, B140 (2005).
[22] M. Reboul, J. Bouvaist, *Rev. Aluminium*, January 1980, p. 41.
[23] T. Shibata, *Corrosion* 52, 813 (1996).
[24] P. J. Laycock, R. A. Cottis, P. A. Scarf, *J. Electrochem. Soc.* 137, 64 (1090).
[25] H. H. Uhlig, R. W. Revie, *Corrosion and Corrosion control*, 3rd edition, John Wiley, New York (1985) p. 75-76.
[26] J. Horvath, H. H. Uhlig, *J. Electrochem. Soc.* 115, 791 (1968).

© 2007, first edition, EPFL Press

[27] E. A. Lislovs, A. P. Bond, *J. Electrochem. Soc.* 122, 719 (1975).
[28] M. Da Cunha Belo, B. Rondot, J. Bergner, J. Montuelle, *Mem. sci. rev. métallurg*. 70, 725 (1973).
[29] H. Kaesche, *Die Korrosion der Metalle*, 3rd edition, Springer Verlag (1990) p. 306.
[30] S. Smialovska, *Pitting Corrosion of Metals*, NACE, Houston, Texas (1986) p. 225.
[31] G. Frankel, *J. Electrochem. Soc.* 145, 2186 (1998).
[32] H. Böhni, T. Suter, *Electrochim. Acta* 47, 191 (2001).
[33] M. M. Lohrengel, A. Moehring, M. Pilaski, *Electrochim. Acta* 47, 137 (2001).
[34] H. Böhni, T. Suter, A. Schreyer, *Electrochim. Acta* 40, 1361 (1995).
[35] F. Hunkeler, H. Böhni, *Material und Technik* 13, 77 (1985).
[36] P. Heimgartner, H. Böhni, *Corrosion,* NACE 41, 715 (1985).
[37] D. Landolt, O. Zinger, P.-F. Chauvy, *Electrochim. Acta* 48, 3185 (2003).

© 2007, first edition, EPFL Press

CHAPTER 8

OXIDATION AND ATMOSPHERIC CORROSION

8.1 LOW-TEMPERATURE OXIDATION

8.1.1 Reaction of metals with oxygen

Most metals are thermodynamically unstable in contact with oxygen; when exposed to air they react spontaneously to form an oxide. The reaction of a metal or semiconductor with oxygen that takes place in the absence of an electrolyte at temperatures where solid-state diffusion plays a negligible role is referred to as ***low-temperature oxidation***. During low temperature oxidation, the rate of oxide growth decreases strongly with increasing film thickness, and becomes practically zero once the oxide thickness reaches a few nanometers. As a result, low-temperature oxidation involves the transformation of an extremely small amount of metal and generally poses no major problems in terms of corrosion. On the other hand, the presence of a thin oxide film greatly affects the surface properties of metals such as adhesion, friction and wear. Low temperature oxidation of silicon plays an important role in the fabrication of integrated circuits, where silicon oxide serves as a dielectric.

Low temperature oxidation is typically observed at temperatures well below 300 °C. As the temperature rises, solid state diffusion becomes increasingly important and eventually controls the rate of oxide growth. Metal oxidation under these conditions is referred to as ***high-temperature oxidation*** and will be discussed in Chapter 9. High temperature oxidation generally leads to much higher film thicknesses than low temperature oxidation. For a given temperature, the contribution of diffusion to oxide growth varies with the melting temperature of the oxide, because the value of the solid state diffusion coefficient depends on the binding energy of a material and hence on its melting temperature. As a consequence, there is not a single well-defined temperature that separates high-temperature from low-temperature oxidation behavior of materials.

Atmospheric corrosion refers to the reaction of a metal with atmospheric oxygen when, at the surface, humidity and pollutants form an electrolyte. In contrast to low-temperature oxidation, the oxides and other corrosion products thus created form mostly non-compact scales that do not protect the metal from further corrosion.

© 2007, first edition, EPFL Press

Atmospheric corrosion is an electrochemical process and its rate is governed the anodic and cathodic partial reactions taking place at the metal-electrolyte and oxide-electrolyte interfaces. The electrochemical mechanism of atmospheric corrosion resembles that of corrosion in aqueous solution, with two important differences: firstly, the corrosion products stay on the surface, rather than being swept away by the electrolyte and, secondly, the electrolyte periodically evaporates during dry periods, then reforms during wet periods, when the metal is exposed to high humidity.

8.1.2 Experimental study of low-temperature oxidation

The first task in any experimental study of low temperature corrosion is to prepare a metal surface free of oxide and adsorbed species. For this the sample must be placed in an ultra- high vacuum (UHV) chamber. One then eliminates the natural oxide layer and adsorbed impurities by ion sputtering. Because ion bombardment disturbs the uppermost atomic layers of the metal (Chapter 3), a thermal treatment is sometimes applied in order to reestablish the original surface structure. Once a clean metal surface is available one introduces a known amount of oxygen into the UHV chamber by setting its partial pressure, and one follows the evolution of the reaction using a suitable method. Surface sensitive methods for the study of adsorption and thin film growth include surface analysis by AES, XPS, SIMS, optical methods, in particular ellipsometry, or mass sensitive methods such as the quartz microbalance.

Figure 8.1 shows typical results from a study of iron oxidation, carried out using an optical method [1]. Here we see how the average thickness of the oxide layer changes as a function of the exposure time at different oxygen partial pressures. The curves include three zones that differ in slope. A zone of initially small slope is followed by one of higher slope, and then by an approximately linear region indicating logarithmic growth of the oxide film. After prolonged time, this last stage results in an oxide thickness of about 3 nm. The initial phase is longest at low partial pressures of oxygen and it does not appear at all at the highest partial pressures.

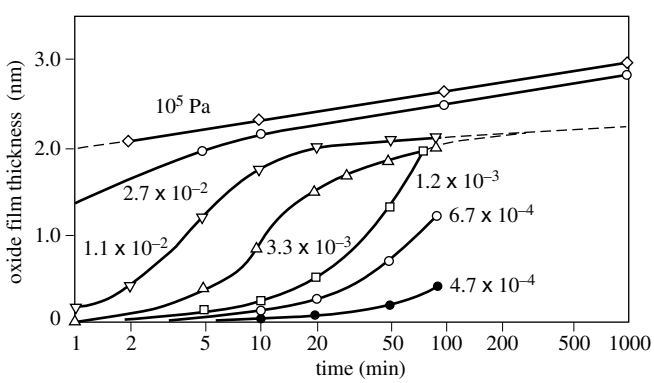


Figure 8.1 Room temperature oxidation of Fe at low oxygen pressure: oxide film growth as a function of exposure time at different oxygen partial pressures [1].

© 2007, first edition, EPFL Press

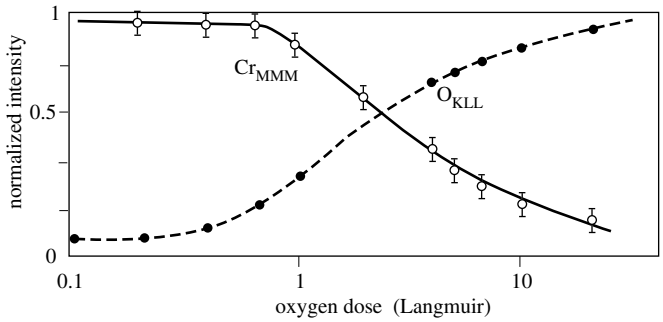


Figure 8.2 Auger electron peak intensities for Cr and O as a function of oxygen dose measured under UHV conditions [2].

The determining parameter in these curves is the ***oxygen dose***, defined as the exposure time multiplied by the partial pressure. Its unit is the ***Langmuir*** L. One Langmuir represents one second of exposure to a partial pressure of 133 Pa (1 torr): 1 L = 133 Pa s.

Figure 8.2 shows the relative variations of the Cr_{MMM} and O_{KLL} Auger peak intensities as a function of oxygen dose during the oxidation of chromium [2]. The intensities measured at different partial pressures form a single curve, thus confirming that the extent of oxidation depends on the oxygen dose, rather than on the oxidation time.

The zones (I), (II) and (III) that make up the oxidation curves shown in Figures 8.1 and 8.2 correspond to the three reaction steps represented schematically in Figure 8.3:

- chemisorption of oxygen;
- lateral growth of oxide islands;
- growth (in thickness) of the oxide film.

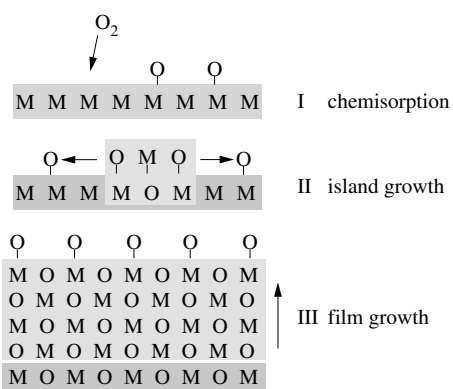


Figure 8.3 Reaction steps between metal surface and oxygen leading to oxide film formation: (I) chemisorption; (II) island growth; and (III) film growth.

© 2007, first edition, EPFL Press

Oxygen molecules collide with the metal surface and are transformed into chemisorbed oxygen atoms:

$$O_2 \rightarrow 2\,O_{ads} \tag{8.1}$$

According to equation (3.44), the number of moles of oxygen that strike the metal surface per second is given by:

$$N_{O_2} = \frac{P_{O_2}}{\left(2\pi M_{O_2} RT\right)^{1/2}} \tag{8.2}$$

where M_{O_2} represents the molecular mass of oxygen. The surface becomes covered with oxygen at a rate that is proportional to the oxygen flux:

$$\frac{d\theta_0}{dt} = \frac{2\,s_k\,N_L\,N_{O_2}}{w_t} \tag{8.3}$$

In this equation, s_k refers to the condensation coefficient (c.f., Chap. 3) and N_L is Avogadro's number. By definition, the surface coverage with oxygen is equal to $\theta_O = w_O/w_t$, where w_O represents the number of sites occupied by the oxygen atoms, and w_t is the total number of adsorption sites.

If the oxygen were adsorbed in its molecular form onto random surface sites, the Langmuir isotherm would yield a condensation coefficient of $s_k = 1 - \theta_O$. However, during adsorption molecular oxygen dissociates into two atoms whose relative position varies in function of the surface crystal structure. It is therefore necessary to modify the condensation coefficient, which is written in the following manner: $s_k = 1 - \sigma\,\theta_O$, where σ represents a dimensionless proportionality factor, with a value between 2 and 4 [3]. The integration of (8.3) gives:

$$\theta_0 = \frac{1}{\sigma}\left[1 - \exp\left(-\frac{2\,N_L N_{O_2}\sigma\,t}{w_t}\right)\right] \tag{8.4}$$

With (8.2), and by introducing the oxygen dose $\vartheta_{O2} = P_{O2}\,t$:

$$\theta_0 = \frac{1}{\sigma}\left[1 - \exp\left(-\frac{2\,\sigma\,N_L\,\vartheta_{O_2}}{w_t\left(2\,\pi\,M_{O_2}\,RT\right)^{1/2}}\right)\right] \tag{8.5}$$

This equation describes the variation of coverage as a function of the oxygen dose. The maximum surface coverage is equal to $\theta_{O,max} = 1/\sigma$.

Island formation

The adsorbed oxygen atoms have a strong affinity for the valence electrons of the metal and form chemical bonds that have partial ionic character. They thus acquire a slightly negative charge and become mutually repulsive. The system can then attain a

© 2007, first edition, EPFL Press

more favorable energy configuration if certain oxygen atoms exchange their position with metal atoms. This phenomenon leads to the formation of an oxide nucleus of monolayer thickness. To grow into a stable island the nucleus must exceed a critical size. The formation of a stable island therefore necessitates several place exchanges. This explains the observation of an incubation time corresponding to zone (I) discussed above.

The lateral growth of islands, and their subsequent coalescence, results in the formation of an oxide layer. Experimental studies with nickel suggest that the oxide surface coverage increases with time according to [3]:

$$\theta_{ox} = 1 - \exp(-k_x t^2) \tag{8.6}$$

Here, k_x is a constant. Equation (8.6) has been explained by different models, assuming that either the surface diffusion of chemisorbed oxygen towards the islands or the flux of molecular oxygen that strikes the metal surface is rate limiting.

Film growth

Stages (I) and (II) are only observable for very low oxygen doses. As soon as the partial pressure becomes elevated, for example equivalent to that of the atmospheric pressure, the oxide layer covers the surface instantly and only the stage (III) remains measurable. The rate of oxide growth then determines the reaction kinetics.

Figure 8.4 summarizes the reactions that contribute to the growth of a compact oxide film. At the oxide-gas interface, the oxygen is reduced to O^{2-} by accepting two electrons. At the metal-oxide interface, the metal atoms oxidize into cations M^{n+} and thereby liberate n electrons. These electrons have to move across the oxide in order to react with the oxygen. Furthermore, the M^{n+} cations move towards the outside surface of the film, while the O^{2-} anions migrate towards the interior. The slowest of these processes determines the growth rate.

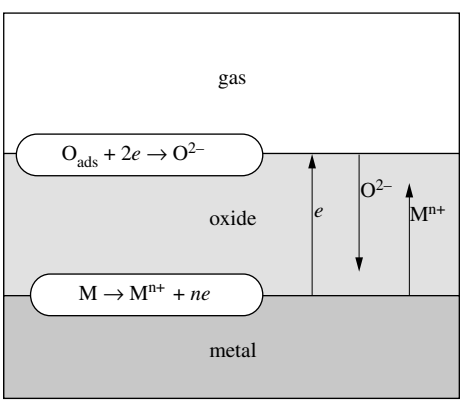


Figure 8.4 Partial reactions during oxide film growth.

© 2007, first edition, EPFL Press

In practice, film growth follows either a logarithmic relation:

$$L \propto \ln t \tag{8.7}$$

or an inverse logarithmic relation:

$$\frac{1}{L} \propto -\ln t \tag{8.8}$$

where L represents the film thickness and t is the oxidation time. In both cases, the growth rate strongly decreases with time to finally reach a negligible value. The final oxide thickness hardly ever exceeds the value of two to three nanometers.

8.1.3 Models of oxide-film growth

Growth limited by electron tunneling

The metal-oxide interface represents an energy barrier to the electrons. The height of the barrier, ΔU_e, corresponds to the difference between the Fermi level of the metal and the conduction band of the oxide. If the energy barrier is not too high electrons can readily surmount it thanks to their thermal energy (Figure 8.5 (a)). They thus move from the metal into the oxide by ***thermal emission***, with a probability W_{th}:

$$W_{th} = \exp\left(-\frac{\Delta U_e}{kT}\right) \tag{8.9}$$

When the energy barrier proves to be much greater than the thermal energy of the electrons ($\Delta U_e >> kT$), the number of electrons that manage to overcome this barrier becomes negligible.

However, another mechanism allows the electrons to cross the interface and the oxide film, provided it is sufficiently thin: ***electron tunneling*** (Fig. 8.5(b)). In this case, the probability that an electron will cross the film is given by W_{tu}:

$$W_{tu} = \exp\left(-\frac{L}{L_t}\right) \tag{8.10}$$

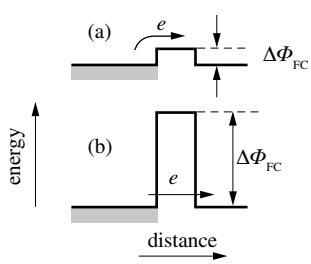


Figure 8.5 Electron emission from a metal in presence of oxide film: (a) thermal emission; (b) emission by electron tunneling.

© 2007, first edition, EPFL Press

$$\frac{1}{L_t} = \frac{4\pi}{h}\left(2\,m_e \Delta U_e\right)^{1/2}$$

Here, h refers to Planck's constant and m_e is the electron mass. The quantity L_t represents the tunneling distance which is on the order of a nanometer. The following relation describes the electron flux N_e (mol m^{-2} s^{-1}) that crosses the film by tunneling [4]:

$$N_e = A' \exp\left(-\frac{L}{L_t}\right) \tag{8.11}$$

$$A' = \frac{N_L \Delta U_e}{2\,\pi\, h\, L^2}$$

In this equation, N_L is Avogadro's number. The growth rate of the film, dL/dt, is proportional to N_e:

$$\frac{dL}{dt} = k' \exp\left(-\frac{L}{L_t}\right) \tag{8.12}$$

$$k' = \frac{A' M}{\rho}$$

where M is the molecular weight and ρ the density of the oxide. By assuming a thickness L_0 at the initial time t_0, the integration of (8.12) yields a relation that describes the variation of the film thickness as a function of the oxidation time:

$$\exp\left(\frac{L}{L_t}\right) - \exp\left(\frac{L_0}{L_t}\right) = \frac{k'}{L_t}\left(t - t_0\right) \tag{8.13}$$

If initially the surface contains no film at all ($y_o = 0$):

$$\exp\left(\frac{L}{L_t}\right) - 1 = \frac{k'}{L_t}\, t \tag{8.14}$$

By taking the logarithm, we get:

$$L = L_t \ln\left(1 + \frac{k'}{L_t}\, t\right) \tag{8.15}$$

For $k't / L_t >> 1$, the thickness of the film increases with the logarithm of the oxidation time:

$$L = L_t \ln\left(\frac{k'}{L_t}\, t\right) \tag{8.16}$$

© 2007, first edition, EPFL Press

This equation is referred to as the ***logarithmic growth law***.

According to the described model [5], the thickness of the oxide film cannot exceed the maximum distance permitted by tunneling, i.e., between 2 and 3 nanometers. Furthermore, in the same manner as electron tunneling, the rate of film growth does not depend on temperature.

Growth limited by high-field ionic conduction

If the electrons can easily overcome the energy barrier at the metal-oxide interface, ion transport in the oxide limits the growth rate.

At ambient temperature, solid state diffusion is too slow to play a significant role. The ions then move essentially by migration, under the influence of the strong electric field present in the oxide film, which typically reaches values of 10^8 to 10^9 V/m. The electric field results from a charge separation induced by chemisorbed oxygen (Figure 8.6).

By assuming that the charges are located along two parallel planes, separated by the distance L (defined by the thickness of the film), the electric field ε_F is equal to:

$$\varepsilon_F = \frac{\Delta\Phi}{L} = \frac{q}{\varepsilon\,\varepsilon_0} \tag{8.17}$$

where $\Delta\Phi$ represents the potential difference between the metal and the outside surface upon which the oxygen is chemisorbed. It is generally assumed that the potential difference at the metal-oxide interface is negligible. In this case, $\Delta\Phi/L$ is equal to the electric field in the oxide. In equation (8.17), q represents the charge density, ε_o is the permittivity constant and ε is the dielectric constant.

In order to simplify the discussion that follows, we assume that only the cations migrate and that, in addition, their concentration c_+ does not vary with the oxide

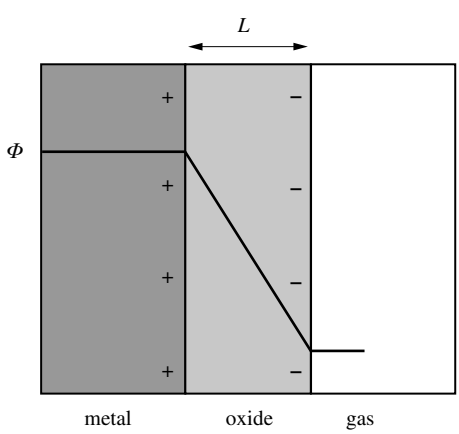


Figure 8.6 Electric field in oxide in contact with oxygen.

© 2007, first edition, EPFL Press

thickness. These conditions correspond to the model presented in Figure 6.22: the electric field lowers the activation energy for the cations migrating towards the outside and increases it for movement in the opposite direction. By neglecting the back migration, the cation flux N_+ is given by:

$$N_+ = 2\,a\,c_+ \nu \exp\left(-\frac{\Delta G^{\#}_{\mathrm{ch}}}{R\,T}\right)\exp\left(\frac{\alpha' z F a\,\varepsilon_{\mathrm{F}}}{R\,T}\right) \tag{8.18}$$

where ν is the vibration frequency, $\Delta G^{\#}_{\mathrm{ch}}$ is the activation energy in the absence of the electric field, z is the charge of the ion, and $2a$ represents the distance between the two equilibrium positions (i.e., the interatomic distance). By taking into account (8.17), equation (8.18) becomes relation (6.22).

During low-temperature oxidation, the potential difference $\Delta\Phi$ is the result of chemical effects. It therefore remains constant during the growth of the oxide layer. By setting:

$$k'' = 2\,a\,c_+ \nu \exp\left(-\frac{\Delta G^{\#}_{\mathrm{ch}}}{R\,T}\right) \tag{8.19}$$

$$L_0 = \frac{\alpha' z F a\,\Delta\Phi}{R\,T} \tag{8.20}$$

we obtain, with (8.17):

$$N_+ = k'' \exp\left(\frac{L_0}{L}\right) \tag{8.21}$$

The growth rate of the oxide is then equal to:

$$\frac{\mathrm{d}L}{\mathrm{d}t} = k_{\mathrm{CM}} \exp\left(\frac{L_0}{L}\right) \tag{8.22}$$

$$k_{\mathrm{CM}} = \frac{M\,k''}{\rho}$$

Integration of (8.22) is not a trivial affair. Cabrera and Mott [6] provided an approximate solution, valid for $L << L_{\mathrm{o}}$:

$$\frac{L_0}{L} = A' - \ln t \tag{8.23}$$

Here, A' designates a constant. Equation (8.23) is identical to (6.32) applicable to potentiostatic anodic oxidation.

© 2007, first edition, EPFL Press

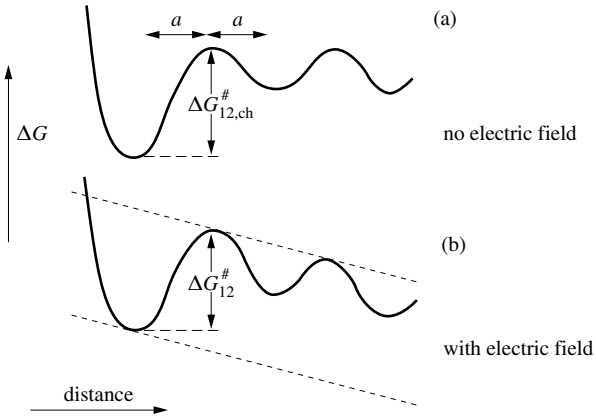


Figure 8.7 Effect of electric field on charge transfer at metal-oxide interface [6].

A more exact integration proposed by Ghez [7] gives:

$$\frac{L_0}{L} = -\ln L_0 k_{CM} - \ln \frac{t + \tau_0}{L^2} \tag{8.24}$$

The adjustable parameter τ_0 has a value close to zero.

Growth limited by cation transfer at the metal-oxide interface

The previous model assumes that the migration of the cations through the film represents the limiting factor and thus corresponds to the high field conduction model for anodic film growth discussed in Chapter 6. Other models can be developed, by considering the charge transfer across one of the interfaces to be rate limiting. Such a model, originally proposed by Cabrera and Mott, is schematically presented in Figure 8.7, assuming the charge transfer at the metal-oxide interface to be rate limiting.

A constant electric field with a value $\Delta\Phi/L$ exists in the oxide that lowers the activation energy $\Delta G^{\#}_{12}$ of the charge transfer reaction at the metal-oxide interface:

$$\Delta G^{\#}_{12} = \Delta G^{\#}_{12,ch} - \frac{\alpha' z F \Delta\Phi}{R T L} \tag{8.25}$$

Mathematically, this model is identical to the previous one and the growth rate is given by equation (8.22), where $\Delta G^{\#}_{ch}$is to be replaced by $\Delta G^{\#}_{12,ch}$. We conclude that the same type of growth law is predicted regardless of whether high field conduction in the film or interfacial charge transfer is rate controlling.

Other models

Two criteria can be used to distinguish oxide growth controlled by the tunneling and by high field conduction: the growth law and the variation of growth rate with

© 2007, first edition, EPFL Press

temperature. Film growth controlled by electron tunneling yields a direct logarithmic growth law independent of temperature. Film growth controlled by high field conduction yields an inverse logarithmic growth law (equation 8.23) and an increase of growth rate with temperature due to the higher ionic mobility. And yet, certain systems show direct logarithmic growth behavior together with a significant effect of temperature.

This apparent contradiction has given rise to a number of alternative models for low-temperature oxidation [8,9]. By assuming, for example, that the activation energy of ion hopping in the high field conduction model is a linear function of the film thickness [8],

$$\Delta G_s^{\#} = \Delta G_{s,o}^{\#} + \lambda L \tag{8.26}$$

one can derive an expression that is mathematically identical to equation (8.21), which represents the direct logarithmic growth. Thus one can reconcile a logarithmic growth law with a temperature dependence of growth rate. However the physical basis of equation (8.26) remains somewhat obscure.

8.1.4 Transport mechanisms

Contribution of anions

In writing equation (8.18) we assumed that only cations migrate. However, in many oxides, anions also contribute to the ionic flux. Sometimes, they even dominate the growth mechanism. Table 8.8 shows the transport numbers of cations, t_+, and anions, t_-, for a number of oxides. The transport number indicates the fraction of the ionic current carried by a given species. For binary oxides that contain only one type of cations and anions: $t_+ + t_- = 1$.

Table 8.8 Transport number of cations and anions for different oxides [8].

Group	Oxide	t_+	t_-
Ib	Cu_2O	1	0
IIIa	Al_2O_3	0.6 – 0.4	0.4 – 0.6
IVb	ZrO_2	< 0.05	> 0.95
IVa	SiO_2	< 0.05	> 0.95
Vb	V_2O_5	0.28	0.72
Vb	Ta_2O_5	0.24	0.76
VIb	WO_3	0.37	0.63

In general, cationic charge transport dominates in crystalline oxides (e.g., Cu_2O, Fe_2O_3), whereas in amorphous oxides (e.g., Ta_2O_5 , SiO_2 , B_2O_3, WO_3) the transport of anions is often predominant.

The model of Cabrera and Mott is also applicable to growth due to the transport of anions. However, under mixed conditions, a more rigorous treatment of the ionic

© 2007, first edition, EPFL Press

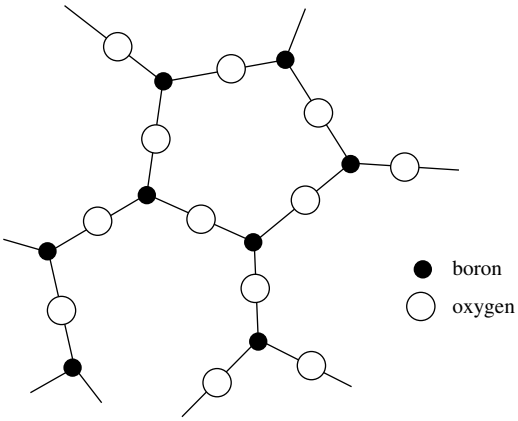


Figure 8.9 Network structure of B_2O_3.

fluxes is necessary taking into account all mobile ionic and electronic species. Because of their complexity, theoretical predictions from such models are generally difficult to verify.

Influence of the oxide structure

The high-field conduction model assumes that the potential energy of the ions is a periodic function related to the spatial arrangement of the atoms. It applies therefore best to crystalline oxides. On the other hand, the model is quite general because it specifies neither the nature of the mobile ions nor the ionic transport mechanism on a microscopic scale. In a crystalline oxide, ionic transport typically involves vacancies or interstitial ions.

Amorphous oxides form less dense networks, in which the cations are attached to the oxygen anions by partially covalent bonds. Figure 8.9 shows a schematic representation of this kind of structure, taking boron oxide as an example. When such a film grows by anion migration, the atom chains break periodically and new bonds are formed. The process depends on the organization of the chains, and in particular on the distance that separates them.

Transport mechanisms in amorphous oxides therefore are more complicated than in crystalline solids, and only partially elucidated. For example, if the structure of an amorphous oxide varies with the film thickness, the transport properties would also depend on thickness. This could offer an explanation for the hypothesis (8.26) made to justify a direct logarithmic growth law. In the same fashion, alloying elements are able to modify the structure of the growing oxide film and thereby influence the transport mechanism and the oxidation rate.

By extension, the model of Cabrera and Mott is often used to describe the growth of nanocrystalline or amorphous oxides by assuming that the potential energy of the mobile ions conserves, on average, a certain periodicity.

© 2007, first edition, EPFL Press

8.2 ATMOSPHERIC CORROSION OF STEEL

8.2.1 Corrosivity of the atmosphere

Chemical composition of ambient air

Ambient air is composed of approximately 77% N_2, 21% O_2, 1% Ar, 1–2% H_2O, 0.03% CO_2 (at %) and traces of rare gases (He, Kr, etc.). In addition, it contains pollutants such as SO_2 and NO_x, chloride ions and airborne particles.

From the corrosion point of view, sulfur dioxide is the most critical pollutant. It originates primarily from the combustion of fossil fuels such as oil and coal in domestic heating systems and industrial power plants.

The formula NO_x refers to a whole family of nitrogen oxides. They are formed in high-temperature combustion, particularly in automobile motors. In terms of corrosion, NO_x plays a lesser role than either SO_2 or chloride ions.

Chlorides are found as air pollutants mostly in maritime regions, where they are carried by wind-borne water droplets (aerosols). The salting of roads in winter represents another source of chloride that is of particular importance for the corrosion of vehicles and highway bridges.

Airborne particles can have both natural and man-made origins. Soot particles from domestic heating, power plants, waste incineration or diesel engines are the result of an incomplete combustion. They can adsorb sulfur species that make them particularly corrosive. Natural sources of airborne particles include volcanoes, trees, and wind swept dry earth.

The ***corrosivity*** of an atmosphere depends mostly on three parameters:

- the relative humidity;
- the SO_2 concentration:
- the chloride concentration.

Humidity

The ***absolute humidity*** (g/m^3) indicates the mass of water contained in a cubic meter of gas, whereas the ***relative humidity*** represents the ratio between the actual vapor pressure and the saturation pressure. The ***saturation pressure*** is that value of vapor pressure, which leads to condensation. The relative humidity is expressed as a percentage. For an atmosphere of given absolute humidity, the value of its relative humidity varies with the temperature (Figure 8.10).

The relative humidity is the most critical parameter for atmospheric corrosion, because it determines whether condensation can take place. At the metal surface, condensed water forms an electrolyte with the salts deposited from pollutants and thus permits electrochemical reactions to take place. In principle, condensation of water occurs when the relative humidity reaches 100%. However, in practice it takes place often at lower values of relative humidity:

- if the temperature of the metal is below that of the ambient air;
- if the surface contains a salt;
- if the surface is porous.

© 2007, first edition, EPFL Press

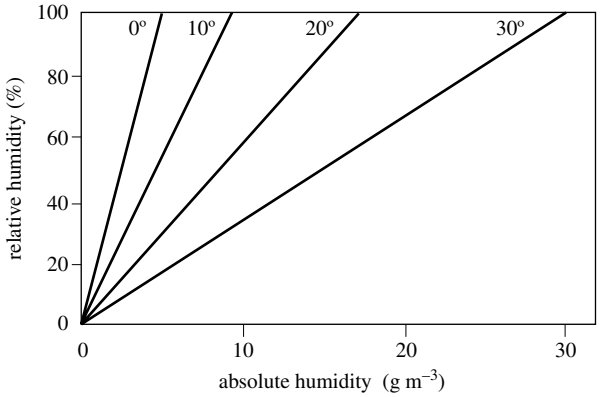


Figure 8.10 Relative humidity as a function of absolute humidity for different temperatures in °C.

The temperature of a metal exposed to the atmosphere is sometimes lower than that of the air itself. A well know example is the condensation of water on the outside wall of a poorly insulated cold water pipe. The same phenomenon may occur when a car that was previously exposed to cold air is parked in a heated garage.

Salts present on the surface due to pollution reduce the saturation vapor pressure. They also increase the ionic conductivity of the electrolyte that forms on the metal surface, thus favoring corrosion even more. Table 8.11 gives the relative humidity in equilibrium with solutions of different salts at their saturation concentration. The given values demonstrate that in presence of certain salts condensation can take place at quite low values of relative humidity. In the laboratory, one uses this effect to control the relative humidity during atmospheric corrosion experiments: one exposes the sample to air in a chamber that contains a salt solution of known concentration. The saturation vapor pressure then determines the humidity of the atmosphere for any given temperature.

Table 8.11 Relative humidity (in %) leading to condensation in equilibrium with saturated salt solutions at 20 °C.

$CuSO_4$	$FeSO_4$	$ZnSO_4$	NaCl	$FeCl_2$	$NiCl_2$	$FeCl_3$	$CaCl_2$	$ZnCl_2$
98	92	90	76	56	54	47	32	10

The presence of porous corrosion products or particles on a metal surface can also facilitate condensation. Indeed, according to the Kelvin equation (3.15), the vapor pressure P_c in equilibrium with water contained in a pore of radius r is lower than the saturation pressure the P_{sat} for a flat surface.

$$\ln \frac{P_c}{P_{sat}} = -\frac{2\gamma M}{\rho R T r} \tag{8.27}$$

© 2007, first edition, EPFL Press

Here, γ represents the surface tension, M is the molecular weight, and ρ is the density of water.

Time of wetness

The relative humidity of ambient air varies as a function of temperature and of climatic conditions such as sun, rain or wind. The corrosivity of the atmosphere, and therefore the rate of corrosion, varies accordingly. We define the ***time of wetness*** as the period of time during which the relative humidity exceeds 80% at a temperature above 0 °C [10]. The time of wetness thus takes into account, in a general manner, the different factors that influence condensation.

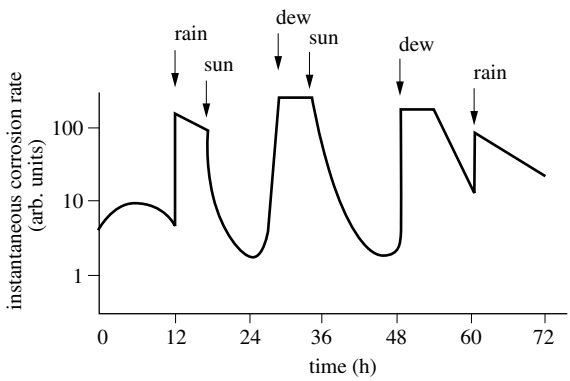


Figure 8.12 Variation of instantaneous corrosion rate with weathering conditions [11].

Figure 8.12 highlights the determining influence of the time of wetness for the corrosion of steel. The instantaneous corrosion rate is plotted versus time and climatic changes are indicated. During dry periods, the instantaneous corrosion rate is one to two orders of magnitude lower. As a first approximation, the total thickness of metal lost to corrosion L is therefore given by the sum of the times of wetness $\tau_{h,k}$ multiplied by v_k, the average rate of corrosion during these periods.

$$L = \Sigma\, v_k \tau_{h,k} \tag{8.28}$$

Figure 8.13 and 8.14 further emphasize the role of the relative humidity. The Figure 8.13 indicates that the corrosion rate of steel increases markedly as soon as the relative humidity exceeds 80%. In order to accelerate corrosion testing, all samples had been previously subjected to a polluted urban atmosphere for a month. Figure 8.14 shows the amount of corroded material as a function of the relative humidity for steel samples that were exposed for 55 days to an atmosphere containing 0.01% SO_2. In this case, corrosion becomes important already at a relative humidity above 60%. The difference with respect to the results of the previous figure is due to the higher levels of SO_2 which favor the presence of sulfate on the surface.

© 2007, first edition, EPFL Press

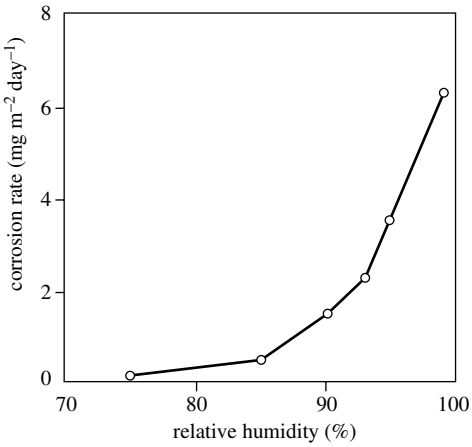


Figure 8.13 Effect of relative humidity on corrosion rate of steel [11].

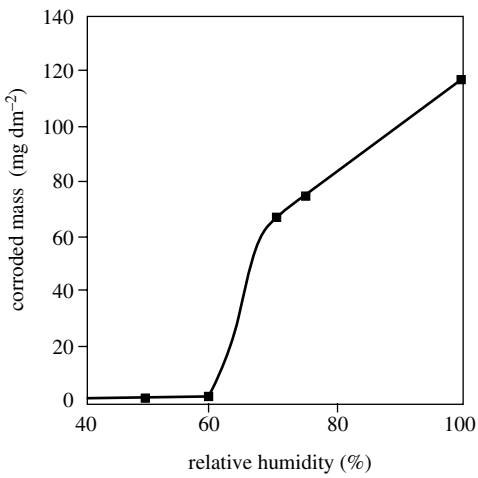


Figure 8.14 Variation of the corroded mass as a function of the relative humidity during laboratory tests, of 55-day duration, in the presence of 0.01% SO_2 [12].

SO_2 concentration

Worldwide, some 100 millions tons of SO_2 are emitted annually into the atmosphere by combustion of oil and coal. The SO_2 concentration in the air, for a given locality, varies greatly depending on its proximity to sources of pollution and prevailing climatic conditions, notably the wind direction. It is expressed in ppm, in mg m^{-3} or in mg m^{-2} day^{-1}. The last set of units refers to the total quantity of sulfur oxides absorbed onto an alkaline filter according to a standard procedure. The SO_2 concentration of a non-polluted atmosphere (rural atmosphere) is typically under 10 μg m^{-3} and that of a moderately polluted atmosphere (urban environments) reaches

© 2007, first edition, EPFL Press

10–100 μg m^{-3}. In an industrial, heavily polluted atmosphere, it can exceed 100 μg m^{-3}.

Atmospheric sulfur dioxide reacts with water to form sulfurous acid:

$$SO_2 + H_2O \rightarrow H_2SO_3 \tag{8.29}$$

In the presence of oxygen and catalytic species, in particular ferrous and ferric ions, it is transformed into sulfuric acid:

$$SO_2 + H_2O + 1/2\ O_2 \rightarrow H_2SO_4 \tag{8.30}$$

The reaction with iron produces ferrous sulfate:

$$4\ H_2SO_4 + 4\ Fe + 2\ O_2 \rightarrow 4\ FeSO_4 + 4\ H_2O \tag{8.31}$$

Reactions (8.29) and (8.30) are responsible for the phenomenon of "acid rain": in polluted areas, the pH of rainwater and of fog can reach relatively low values, typically between 4 and 5. At the surface of exposed steel, the pH remains higher (5 to 7), because the reaction (8.31) consumes sulfuric acid. This explains the fact that proton reduction ($2\ H^+ + 2e \rightarrow H_2$) is less important for the atmospheric corrosion of steel than oxygen reduction.

Water condensed at the surface of the metal, forms an electrolyte whose conductivity increases with the concentration of the sulfate ions. This facilitates the formation of electrochemical cells and greatly accelerates the rate of corrosion. Figure 8.15 illustrates the role of SO_2 for the corrosion of iron. It shows the results of laboratory tests carried out by exposing samples to an atmosphere of 95 % relative humidity, with and without SO_2. Curve (a) presents the corroded mass as a function of exposition time observed in an SO_2-free atmosphere. Curve (b) shows the results of an identical test carried out in the presence of SO_2: the weight loss is much more significant. Curve (c) corresponds to a hybrid experiment. The samples were first exposed for a short period of time (4 hours) to an atmosphere containing SO_2 (identical to that of curve (b)), and then to an SO_2-free atmosphere (corresponding to the experiment indicated as curve (a)). Interestingly, despite the absence of SO_2, the result is almost identical to that of curve (b). Curve (d) was produced using an analogous procedure except that the samples were washed with water after the 4 hours exposure to the polluted atmosphere. The corroded mass now is comparable to that observed during the experiment corresponding to curve (a).

The results presented in Figure 8.15 allow us to draw the following conclusions:

- The presence of SO_2 in the atmosphere greatly increases the rate of corrosion (by comparison of tests (a) and (b));
- A sulfate-contaminated surface corrodes rapidly, even if the atmosphere does not contain SO_2 (test (c));
- By eliminating the contaminant deposits from the surface one can re-establish a slow corrosion rate (test (d)).

© 2007, first edition, EPFL Press

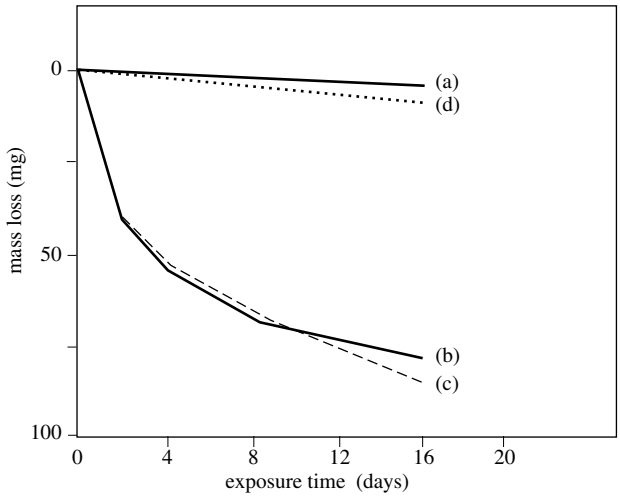


Figure 8.15 Weight loss measured on iron samples exposed to air with relative humidiy of 95% under laboratory conditions: (a) in the presence of SO_2; and (b) in the absence of SO_2. The curve (c) represents the results obtained by exposing a precorroded sample in the presence of SO_2 (and therefore contaminated by sulfates)to an atmosphere without SO_2. The curve (d) represents the same kind of test as (c), except that the sample had been rinsed with water before the corrosion test [13].

The results of Figure 8.16 further highlight the importance of pollution due to SO_2. The mass of corroded steel is presented as a function of the time of exposure to different atmospheres. Corrosion increases in the order:

Rural atmosphere < urban atmosphere < industrial atmosphere

We also observe that the average corrosion rate (given by the slope of the curves) is not constant. It is highest at the start of the experiments, then it decreases with time, finally reaching a constant value after a period of several years. The initial period, corresponding to a non-steady-state corrosion rate, shortens with increasing degree of atmospheric pollution.

Chloride concentration

In the atmosphere, one finds chloride ions dissolved in suspended water droplets. The chloride content of an atmosphere is measured in mg/m^2 day and corresponds to the amount of chloride absorbed by an aqueous solution under standard conditions over a period of one day. Chloride concentrations as high as 1500 mg/m^{2}day can be reached in marine environments.

In a similar way as sulfates, chlorides in presence of water form an electrolyte on the metal surface, thus creating conditions that favor corrosion. In addition, they render passivation more difficult and they may cause pitting (cf. Chapter 6).

© 2007, first edition, EPFL Press

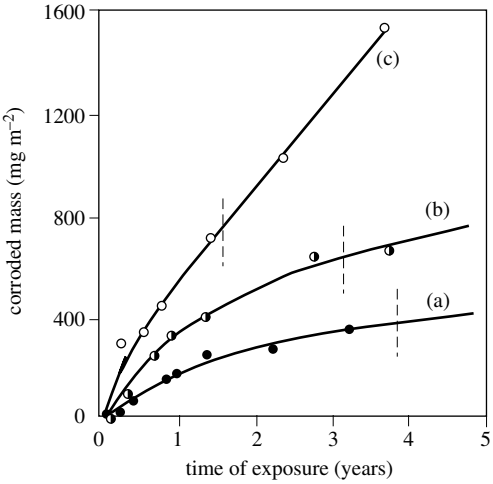


Figure 8.16 Corroded mass of a steel sample (a) exposed to a rural atmosphere; (b) to an urban atmosphere; and to (c) an industrial atmosphere for a period of several years. The dashed lines indicate the time required to reach the steady-state corrosion rate [14].

Classification of atmospheres according to their corrosivity

The time of wetness, which depends on climatic conditions, and the concentrations of SO_2 and Cl^-, have been used as criteria for defining five types of atmosphere which differ in their corrosivity [10]. Table 8.17 indicates the average corrosion rate of steel for each of them. We may recall that during the initial phase, which can last several years, the corrosion rate decreases with the exposure time (Figure 8.16). Average rates of atmospheric corrosion given in tables therefore depend on the total exposure time used in the tests. The data of Table 8.17 confirm a decreasing of corrosion rate with time until a steady value is reached after several years.

The corrosion rates of Table 8.17 refer to outdoor exposure conditions. Inside buildings, corrosion is generally much less pronounced because of the absence of rain, wind and of extreme changes in temperature and humidity.

Table 8.17 Average corrosion rate of steel in different atmospheres [10]

Corrosivity of the atmosphere	Average corrosion rate for the first year μm/year	Average corrosion rate for the first ten years μm/year	Average rate of corrosion at steady state μm/year
very weak	< 1.5	< 0.5	< 0.1
weak	1.5 – 25	0.5 – 5	0.1 – 1.5
average	25 – 50	5 – 12	1.5 – 6
strong	50 – 80	12 – 30	6 – 20
very strong	80 – 200	30 – 100	20 – 90

© 2007, first edition, EPFL Press

8.2.2 Corrosion products

The atmospheric corrosion of steel, commonly referred to as *rusting* of steel, can be represented in a simplified way by the following stoichiometric equation:

$$4\,Fe + 3\,O_2 + 2\,H_2O \rightarrow 4\,FeOOH \tag{8.32}$$

The chemical formula FeOOH designates ***rust***, a hydrated oxide with a stoichiometry intermediate between ferric oxide, Fe_2O_3, and ferric hydroxide, $Fe(OH)_3$. The following stoichiometric equivalences demonstrate this:

$$2\,FeOOH \leftrightarrow Fe_2O_3 \cdot H_2O$$

$$2\,Fe(OH)_3 \leftrightarrow Fe_2O_3 \cdot 3\,H_2O$$

In reality, the reaction of iron with oxygen and water yields, depending on conditions, a multitude of oxides and hydroxides (Table 8.18) and the composition of rust varies accordingly.

Table 8.18 Common oxides and hydroxides of iron.

Formula	Name	Crystal system
α-Fe_2O_3	hematite	trigonal
γ-Fe_2O_3	maghemite	cubic
Fe_3O_4	magnetite	cubic
FeO	wustite	cubic
α-FeOOH	goethite	rhomboedric
β-FeOOH	acaganeite	tetragonal
γ-FeOOH	lepidocrocite	rhomboedric
δ-FeOOH		hexagonal
$Fe(OH)_2$	ferrous hydroxide	
$Fe(OH)_3$	ferric hydroxide	
$Fe(OH)_3 \bullet 4Fe(OH)_2 \bullet FeSO_4 \bullet nH_2O$	type II green rust	

Constituents of rust

The stoichiometry of rust only approximately corresponds to the formula FeOOH. In reality, rust is a complex mixture of different phases (crystalline and amorphous) of the oxides and hydroxides of iron. Most of the iron in rust is in the trivalent oxidation state, but bivalent iron is also present. The chemical composition of rust and its structure vary as a function of the climatic conditions and of the length of exposure to the atmosphere. In addition, it usually differs between the interior and exterior parts of the layer.

The three principal components of rust are:

- α-FeOOH (goethite);
- γ-FeOOH (lepidocrocite);
- Fe_3O_4 (magnetite).

© 2007, first edition, EPFL Press

Rust layers normally comprise a relatively dense internal region (adjacent to the metal) consisting of magnetite and amorphous phases of composition FeOOH and a porous external region consisting mainly of lepidocrocite and goethite.

Magnetite

Magnetite, Fe_3O_4, is a crystalline phase that contains an equal number of Fe^{2+} and Fe^{3+} ions. Its crystal structure is that of an inverse spinel with an oxygen anion lattice containing octahedral and tetrahedral sites that are occupied by metal ions. The Fe^{3+} ions occupy the tetrahedral sites and half of the octahedral sites, whereas Fe^{2+} is found exclusively on octahedral sites. In contrast to FeOOH phases, magnetite conducts electricity. In addition, it is magnetic.

Magnetite can be formed by different reactions:

- the oxidation of $Fe(OH)_2$ or of "green rust" under conditions of lack of oxygen;
- the electrochemical reduction of γ-FeOOH;
- the reaction of iron with water vapor (Schikorr reaction)

The ***Schikorr reaction***, has the following stoichiometry:

$$3\ Fe + 4\ H_2O \rightarrow Fe_3O_4 + 4\ H_2 \tag{8.33}$$

It takes place only at relatively high temperatures, typically 150 °C, and it results in the formation of compact layers of magnetite, that provide good corrosion protection. In practice, the Schikorr reaction plays a role for the corrosion protection of steam condenser tubes made of steel.

In atmospheric corrosion, the magnetite found in rust results either from the electrochemical reduction of lepidocrocite γ-FeOOH or from the oxidation of "green rust". These reactions do not produce compact magnetite films and, as a result, the magnetite formed does not protect the metal. On the contrary, it can actually be harmful; because of its good electric conductivity, magnetite can act as the cathode for oxygen reduction and thus facilitate corrosion [15].

Phases of trivalent iron

The less dense exterior part of the rust layer contains mostly the phases γ-FeOOH and α-FeOOH. The latter has the higher thermodynamic stability of the two but forms more slowly. For this reason, accelerated atmospheric corrosion tests yield lepidocrocite (γ-FeOOH) primarily.

The various phases of rust formed under atmospheric corrosion conditions often contain sulfate or chloride ions. In a marine atmosphere, for example, the β-FeOOH phase contains a small amount of chloride.

The term ***green rust*** designates a variety of intermediate reaction products, generally amorphous, that are found during the atmospheric corrosion of steel in the presence of chlorides and/or sulfates. More precisely, these products form during the transformation of $Fe(OH)_2$ into γ-FeOOH. Green rust consists of mixed ferric and

© 2007, first edition, EPFL Press

ferrous hydroxide compounds with ferrous chloride or sulfate. In the literature one distinguishes type-I green rust, mostly containing chlorides, and type-II green rust which contains sulfate. In its pure state, the latter corresponds to the formula $2Fe(OH)_3 \bullet 4Fe(OH)_2 \bullet FeSO_4 \bullet nH_2O$. In reality, the green rust formed during atmospheric corrosion corresponds rarely to a well-defined stoichiometry, and its sulfate and chloride content varies with the exposure conditions.

Influence of rust on corrosion rate

The presence of rust influences the rate of atmospheric corrosion of steel in two ways. One hand, a porous rust layer increases the surface area exposed to the atmosphere and thus the number of adsorption sites. More pollutant can adsorb onto the surface, and water condenses more easily. On the other hand, rust layers slow down the rate of corrosion by providing a barrier (although imperfect) between the reactive metal surface and the atmosphere and thus reduces access of oxygen.

When a non-corroded steel surface is exposed to the atmosphere, the rate of corrosion is highest. It then diminishes with exposure time as the rust layer builds up. Empirically, the amount of mass corroded as a function of time can be described by the equation:

$$m = k\,t^n \tag{8.34}$$

where m represents the corroded mass, k is a proportionality factor, t is the exposure time, and n is an exponential factor whose value depends on the climatic conditions and the type of steel. Generally, the value of n lies between 0.4 and 0.8.

After an initial period of time, often lasting several years, the corrosion rate reaches a steady state value that only depends on the corrosivity of the environment and on the type of steel. The corroded mass then increases linearly with time, meaning that with further build-up of the rust layer its accelerating and inhibiting effects compensate each other.

8.2.3 Reaction mechanisms

Reaction of a humid iron surface with oxygen

When a polished iron or steel surface is exposed to a humid atmosphere, it becomes promptly covered with a brownish thin film of corrosion products. The reaction is all the more rapid if the surface has been contaminated with ionic species such as sulfates. The reaction mechanisms involved are schematically presented in Figure 8.19.

An aqueous film, some micrometer thick, is present on the metal surface where it forms an electrolyte. Atmospheric oxygen diffuses across the liquid and reacts at the metal surface, while iron is oxidized to ferrous ions that dissolve in the electrolyte (Figure 8.19(a)).

$$Fe \rightarrow Fe^{2+} + 2\,e \tag{8.35}$$

© 2007, first edition, EPFL Press

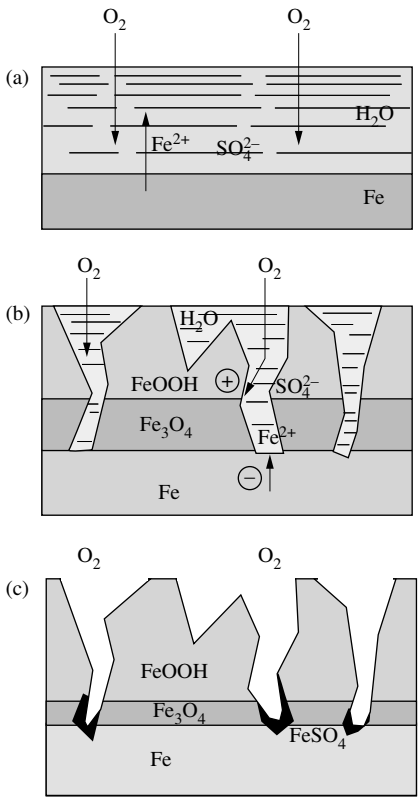


Figure 8.19 A schematic illustrating the mechanism of atmospheric corrosion of steel: (a) shows the inital phase of the corrosion of a metallic surface covered by a thin water film; (b) shows the reactions that take place on a wet surface in the presence of a rust layer; and (c) shows the oxidized state of a dry rusted surface with precipitated iron sulfates.

$$\frac{1}{2}\,O_2 + H_2O + 2\,e \;\rightarrow\; 2\,OH^- \tag{8.36}$$

The ferrous ions react further with oxygen, leading to the precipitation of γ-FeOOH :

$$2\,Fe^{2+} + 3\,H_2O + \frac{1}{2}\,O_2 \;\rightarrow\; 2\,\gamma\text{-FeOOH} + 4\,H^+ \tag{8.37}$$

Equation (8.37) expresses the overall stoichiometry. In fact, the reaction takes place in several steps involving a number of intermediate reaction, in particular "green rust". The γ-FeOOH is not the only reaction product and other crystalline and amorphous phases also form, notably $Fe(OH)_3$.

Under the conditions described in Figure 8.19(a), oxygen transport to the surface limits the corrosion rate of iron. The rate is therefore inversely proportional to the thickness of the liquid film.

© 2007, first edition, EPFL Press

Reactions in the presence of rust

The amount of solid reaction products forming a rust film increases gradually as corrosion progresses. Figures 8.19(b) and (c) represent a steel surface covered with a porous rust layer, exposed to wet (b) and dry (c) conditions, respectively.

On a wet surface, iron dissolves as ferrous ions according to (8.35). This reaction takes place at the base of pores where the metal is in contact with the electrolyte. Two possibilities exist for the cathodic partial reaction:

- reduction of oxygen;
- reduction of γ-FeOOH.

Oxygen reduction following equation (8.36) can take place either at the metal-electrolyte interface situated at the base of a pore or at the surface of the magnetite present in the rust layer. From a diffusion point of view, the magnetite at the pore walls of the rust layer is more accessible than the metal surface at the base of the pore and oxygen reacts preferentially there. As a consequence, an electrochemical cell is established in which the base of the pore becomes the anode and the magnetite the cathode (Figure 8.19(b)).

The reduction of γ-FeOOH involves a reaction with ferrous ions and leads to the formation of magnetite:

$$8\ \gamma\text{-FeOOH} + Fe^{2+} + 2\,e \rightarrow 3\ Fe_3O_4 + 4\ H_2O \tag{8.38}$$

The sum of the anodic (8.35) and cathodic (8.38) partial reactions corresponds to the overall reaction:

$$8\ \gamma\text{-FeOOH} + Fe \rightarrow 3\ Fe_3O_4 + 4\ H_2O \tag{8.39}$$

The γ-FeOOH thus can act as an oxidizing agent during the corrosion of iron in a comparable way with oxygen.

A simple and elegant experiment due to Evans [13] demonstrates that the overall reaction (8.39) indeed contributes to corrosion. The principle is presented in Figure 8.20. Two sheets, one of polished iron and the other of rusted iron, form a sandwich. They are separated with a sheet of filter paper soaked with a sulfate solution. The

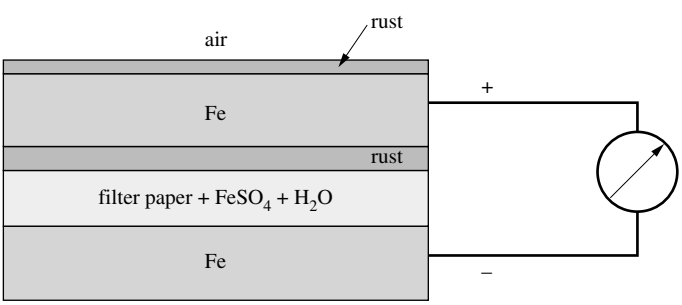


Figure 8.20 Corrosion cell formed between oxidized and non-oxidized iron surfaces in the presence of a sulfate electrolyte [13].

© 2007, first edition, EPFL Press

current between the two metal sheets is then measured as a function of time. At certain intervals, the rusted sheet is turned and the face that was previously exposed to the air is put in contact with the electrolyte. Each change creates a transient current that decreases with time.

The following mechanism explains these observations: An electrochemical cell is established between the metal and the oxide, in which the γ-FeOOH is the cathode and progressively reduces into an oxide of lower valence such as magnetite. The reaction should stop as soon as all γ-FeOOH has been used up. However, by exposing the reduced oxide to the air its original valence state is regenerated each time the cathode sheet is turned:

$$2\,Fe_3O_4 + \frac{1}{2}\,O_2 + 3\,H_2O \rightarrow 6\,\gamma\text{-FeOOH} \tag{8.40}$$

The described experiment demonstrates why ***humidity cycles***, that is to say the periodic switching between wetting and drying of the metal surface, have a determining influence on the rate of atmospheric corrosion of steel.

Instantaneous rate of partial reactions

The results of Figure 8.21 provide further evidence for the participation of two cathodic partial reactions in the atmospheric corrosion of steel. A thin sheet of rusted iron is periodically exposed to humidity in a closed chamber containing oxygen and water vapor. Simultaneously, the rate of iron oxidation (using a magnetic method) and the rate of oxygen reduction (determined volumetrically) are measured. The sample is

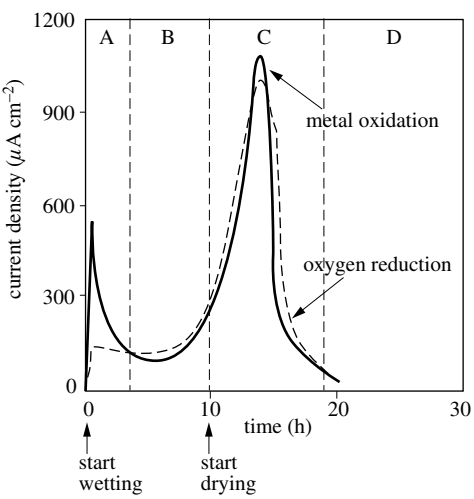


Figure 8.21 Partial currents of Fe oxidation and O_2 reduction on rusted iron exposed to wet-dry cycling: A= wetting; B= wet state; C= drying; and D= dry state (adapted from [16]).

© 2007, first edition, EPFL Press

subjected to humidity cycles by periodically changing the temperature, which controls the relative humidity of the gas.

At the beginning of the wetting period, the iron oxidation rate (expressed as partial current) significantly exceeds that of oxygen reduction. To satisfy the charge balance, a second cathodic reaction is therefore required, namely, the reduction of γ-FeOOH. The anodic partial current is then equal to the sum of the two cathodic partial currents:

$$I_{a,Fe} = I_{c,O2} + I_{c,FeOOH} \tag{8.41}$$

where $I_{a,Fe}$ represents the anodic current of the partial reaction (8.35), whereas $I_{c,O2}$ and $I_{c,FeOOH}$ designate, respectively, the partial current for the reductions of oxygen and of γ-FeOOH.

Reaction (8.38) is not the only possible reduction of γ-FeOOH into magnetite. In the absence of ferrous ions, another reaction, comprised of two steps, can take place:

$$\gamma\text{-FeOOH} + H^+ + e \rightarrow Fe(OH)_{2(surface)} \tag{8.42}$$

$$Fe(OH)_{2(surface)} + 2\ \gamma\text{-FeOOH} \rightarrow Fe_3O_4 + 2\ H_2O \tag{8.43}$$

Here, $Fe(OH)_{2(surface)}$ represents a partial conversion of the surface of γ-FeOOH into ferrous hydroxide. The overall stoichiometry corresponds to

$$3\ \gamma\text{-FeOOH} + H^+ + e \rightarrow Fe_3O_4 + 2\ H_2O \tag{8.44}$$

As the reduction of γ-FeOOH progresses, the magnetite layer thickens. This in turn leads to an increase in the rate of oxygen reduction because of the higher cathodic surface area. Finally, the reduction of γ-FeOOH stops when its content in the rust layer is exhausted. In reality, the presence of oxygen limits the degree of reduction of γ-FeOOH and its transformation is never complete. The exterior part of rust layers therefore always contains some FeOOH.

At the beginning of a drying phase the relative humidity decreases and the condensed water starts to evaporate. Initially, the rate of corrosion increases because the smaller thickness of the liquid layer facilitates the diffusion of oxygen towards the cathode and the decreasing volume increases the concentration of ions. However, once fully evaporated, the electrolytic contact between the anodic and cathodic zones is interrupted and the corrosion cell is stopped. The rate of corrosion of the metal drops sharply at this point. Simultaneously, the rust starts to reoxidize.

When a rusted iron sample dries, its corrosion potential changes from a value of about –400 mV, close to the theoretical value of the Fe/Fe^{2+}electrode, to about +150 mV, typical for the corrosion potential of a passive electrode (Figure 8.22). At such high potentials the ferrous states at the oxide surface can transform into ferric states:

$$2\ Fe(OH)_{2(surface)} + \frac{1}{2}\ O_2 \rightarrow 2\ \gamma\text{-FeOOH} + H_2O \tag{8.45}$$

Drying thus restores the oxidized state of rust, and the cycle can then begin anew.

© 2007, first edition, EPFL Press

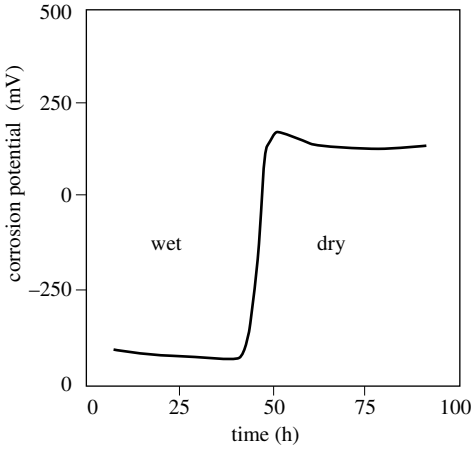


Figure 8.22 Variation of corrosion potential of iron during drying in air [16].

The described mechanism is similar to that corresponding to the equations (8.39) and 8.40) except for two minor aspects. Firstly, the intermediate product $Fe(OH)_{2(surface)}$ is oxidized during drying, rather than the magnetite phase directly, taking into account that crystalline magnetite is chemically rather stable. Secondly, the dissolved ferrous ions do not participate in magnetite formation. Despite some uncertainty concerning the nature of the intermediate products, both mechanisms highlight the fact that rust actively participates in the mechanism of atmospheric corrosion of iron and steel. Most important, they explain the dominating influence of humidity cycles on the corrosion rate.

Effect of sulfates

The formation of a corrosion cell within a pore of the rust layer requires the presence of an electrolyte having sufficient ionic conductivity. This simple fact explains why sulfates and other ionic contaminants, besides facilitating condensation, accelerate the rate of atmospheric corrosion. During drying, as water evaporates, the sulfate anions together with ferrous cations precipitate within the pores, thus forming ferrous sulfate hydrate, $FeSO_4 \bullet nH_2O$ (cf. Figure 8.19(c)). Subsequent wetting fills the pores with water and the resulting dissolution of ferrous sulfate, creates a concentrated electrolyte solution. During the humidity cycles, the electrolyte solution stays inside the pores of the rust layer unless it is washed out by wet cleaning or intensive rainfall.

In conclusion, the atmospheric corrosion of steel involves two distinct surface states, wet and dry, which differ in terms of anodic and cathodic partial reactions taking place. The average rate of corrosion depends above all on the action of humidity cycles and it is accelerated by the presence of sulfates and other anionic contaminants. By forming an electrolyte these permit the creation of corrosion cells between the cathodic and anodic sites in the porous rust that covers the surface.

© 2007, first edition, EPFL Press

8.2.4 Practical aspects

Atmospheric corrosion tests

The described mechanisms of atmospheric corrosion demonstrate the importance of pollutants and of humidity cycles. Unfortunately, the complexity of the underlying phenomena, makes it difficult to predict atmospheric corrosion behavior from accelerated testing. Outdoor exposure tests, which generally require several years, therefore remain popular. Different norms and recommended practices fix in detail the conditions to be applied in atmospheric corrosion tests. The following accelerated test methods are most commonly used in the laboratory:

- salt spray tests;
- environmental chamber tests;
- cyclic immersion tests.

In a ***salt spray test*** the sample is brought in contact with a saline fog in a closed chamber for a certain number of hours. The fog is formed by atomizing a NaCl solution by means of a nozzle. The high corrosivity of the artificial fog simulates that of a marine atmosphere, but the method does not offer a realistic model of more general atmospheric corrosion conditions, because it involves neither the presence of SO_2 nor humidity cycling. In order to accelerate the corrosion, acetic acid or copper chloride is sometimes added to the saline solution.

Environmental test chambers are flexible tools that permit the operator to carry out many different accelerated weathering tests for the evaluation of atmospheric corrosion resistance. Depending on the needs, test procedures include control of gas composition, humidity, temperature and irradiation with artificial sun light. Using computer control, these parameters can be varied almost at will, permitting to impose humidity cycles or a varying pollutant concentration. In order to accelerate the corrosion process, the concentration of polluting gas is usually maintained at a higher level than found in an outdoor atmosphere. A drawback of environmental test chambers is their high cost.

In ***cyclic immersion tests*** the sample is periodically immersed in an electrolyte, for example by attaching it to a partially submerged wheel that slowly turns. With this method, it is easy to simulate the humidity cycles, but there is a major drawback: during the immersion phase, the corrosion products can dissolve, in contrast to what happens normally in true atmospheric corrosion. This procedure therefore simulates atmospheric corrosion only in an imperfect way.

Importance of surface cleaning

According to the electroneutrality condition (4.120), the electrolyte in a pore must contain an equal number of positive and negative charges. Metal dissolution at the bottom of a pore yields ferrous ions, the positive charge of which is compensated by anions, especially sulfates and chlorides. To maintain electroneutrality, the anions migrate towards the base of the pore (Figure 8.19(b)) where they accumulate. This

© 2007, first edition, EPFL Press

increases the ionic conductivity of the electrolyte and thus further facilitates the establishment of corrosion cells. The effect is particularly important at high ionic concentration, in other words, on a highly polluted surface.

Under certain conditions, oxides or hydroxides may precipitate due to a reaction between dissolving metal cations and cathodically generated hydroxyl ions. Because the solubility of ferrous hydroxide is greater than that of ferric hydroxide, the latter precipitates preferentially. It forms by oxidation of ferrous ion on contact with oxygen, and therefore is found mainly in the exterior part of the pores. Under certain conditions, the precipitates can block the pore opening and thereby trap ionic species inside. Such spots containing a particularly high concentration of sulfates are referred to as ***sulfate nests***.

The described processes explain why it is extremely important to well clean a steel surface before applying a protective coating, in particular paints that are water and oxygen permeable as will be discussed in Chapter 12. A simple rinsing with water does not adequately remove the sulfates trapped in the rust layer. As a consequence of their presence corrosion will take place at the metal-coating interface, leading to blistering and flaking of the paint. In fact, the cleanliness of the metal surface often has a greater influence on the lifetime of a coating than the intrinsic quality of the paint used. To clean steel surfaces, before applying a paint, one commonly uses ***sandblasting***. This process consists of projecting a powerful stream of sand, dry or humid, which strips the rust from a surface by abrasion.

Weathering steel

Low alloy steels, also called weathering steels, contain small amounts of copper, chrome, nickel, phosphorus, silicon, and magnesium (< 1%, typically). Their resistance to atmospheric corrosion generally exceeds that of carbon steel. Indeed, when exposed to environments that are not too strongly polluted a dark brown patina forms over some years that slows down the corrosion rate. On buildings this natural layer can thus replace a paint coating. For this reason, alloyed steels find numerous applications in architecture.

The patina, which can be several micrometer thick, has a chemical composition similar to that of rust found on carbon steel. However, it exhibits a different morphology, being less porous. As a consequence, the rate of corrosion of weathering steels is typically 3 to 5 times less than that of ordinary steels. Figure 8.23 illustrates this behavior. Three types of steel were exposed to a polluted atmosphere: a carbon steel, a steel containing a small amount of copper (0.23%) and a low alloy steel (0.38% Mn, 0.93% Cr, 0.09% P, 0.56% Si, 0.44% Cu). The results confirm that the corrosion rate of the steels forming a patina is slowest.

The reaction mechanisms by which the alloy elements influence the rate of corrosion of weathering steel are not precisely known. While the presence of alloy elements improves the corrosion resistance of a carbon steel subjected to humidity cycles, during a simple immersion test weathering steel and carbon steel yield

© 2007, first edition, EPFL Press

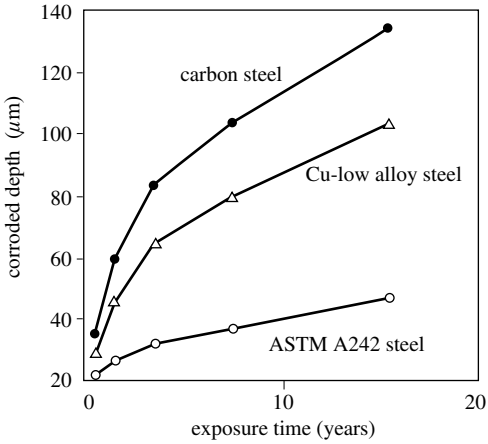


Figure 8.23 Atmospheric corrosion of different types of steel (carbon steel, Cu-low alloy steel, and ASTM A242 steel) in a polluted environment [17].

comparable results.The beneficial effect of adding copper has been attributed to its ability to form basic sulfates of low solubility. By precipitating in the pores they are thought to diminish the corrosive effect of dissolved sulfates. In the same way one can explain the beneficial effect of nickel. The effect of phosphorus may be related to the precipitation of poorly soluble phosphates formed by an oxidation process.

From a practical point of view, two phenomena are of particular importance. Firstly, the protective patinas do not form in strongly polluted atmospheres, for example near a chimney. Under these conditions, weathering steel corrodes just like carbon steel. Indeed, in an environment with an elevated concentration of SO_2, the pH of the surface of the steel drops to low values, thus creating acid conditions which prevent the basic sulfates from precipitating in the pores. Secondly, when rainwater runs off a weathering steel plate it carries with it corrosion products contained in the patina. These products can cause ugly stains on concrete walls situated downstream from the steel element. Proper design is needed to avoid such staining.

8.3 ATMOSPHERIC CORROSION OF NON-FERROUS METALS

Among the non-ferrous metals, zinc, aluminum, and copper are most often employed in outdoor applications such as buildings, installations and transportation systems. Table 8.24 indicates the average corrosion rates of these metals in different types of atmosphere. Compared to steel, these rates are rather small. The values given in parentheses indicate the maximum depth of pits (in μm) after twenty years of exposure time.

© 2007, first edition, EPFL Press

Table 8.24 Rate of atmospheric corrosion of different metals in μm/year. In parenthesis maximum depth of pits after 20 years [18].

Atmosphere	Zinc	Copper	Aluminum
rural	0.2 – 2	< 1	< 0.1 (10 – 55)
urban or industrial	2 – 16	1 – 3	< 1 (100 –190)
marine	0.5 – 8	1 – 2	0.4 – 0.6 (85 – 260)

Atmospheric corrosion of zinc

The atmospheric corrosion rate of zinc generally varies little with time. This indicates that the corrosion products formed at the surface do not markedly influence the reaction rate. On the other hand, the relative humidity and the presence of atmospheric pollutants play a crucial role. Figure 8.25 shows the influence of the SO_2 concentration in the atmosphere. The corrosion rate is plotted versus time of exposure in a heavily polluted urban environment, namely Berlin between 1940 and 1943. The variation in relative humidity and in SO_2 concentration during the same period is also shown. The higher SO_2 concentration observed in winter is due to household heating by burning of sulfur containing coal. Over the different seasons, the corrosion rate closely follows the SO_2 concentration in the atmosphere; both values are highest in winter.

The atmospheric corrosion of zinc involves the following partial reactions:

$$Zn \rightarrow Zn^{2+} + 2\,e \tag{8.46}$$

$$\frac{1}{2}O_2 + H_2O + 2\,e \rightarrow 2\,OH^- \tag{8.36}$$

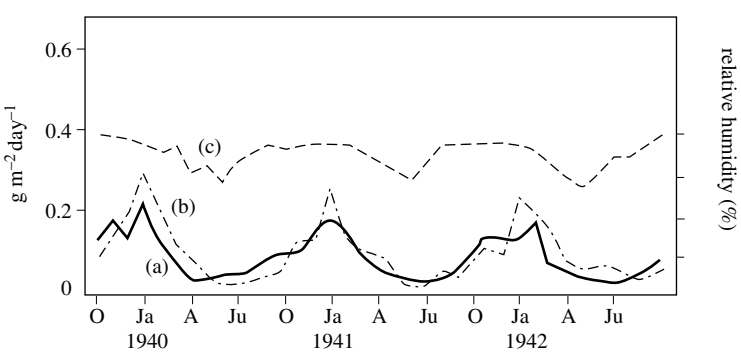


Figure 8.25 Effect of air pollution on the corrosion rate of zinc: (a) monthly variation of the corrosion rate over a period of three years; (b) the atmospheric concentration of SO_2; and (c) the value of the relative humidity [19].

© 2007, first edition, EPFL Press

which corresponds to the overall reaction:

$$Zn + \frac{1}{2} O_2 + H_2O \rightarrow Zn(OH)_2 \tag{8.47}$$

Zinc hydroxide $Zn(OH)_2$ precipitates on the surface. In addition, the oxide ZnO is found among the corrosion products, as well as a number of salts of variable stoichiometry, mostly basic and containing hydroxyl groups. Table 8.26 lists the principle corrosion products of zinc together with those of aluminum, and copper.

Table 8.26 Principle corrosion products of zinc, copper and aluminum.

Metal	Corrosion products
Zn	ZnO, $Zn(OH)_2$, $Zn(OH)_x(CO_3)_y$, $Zn(OH)_x(SO_4)_y$, $Zn(OH)_xCl_y$
Cu	Cu_2O, CuO $Cu(OH)_x(SO_4)_y$, $Cu(OH)_xCl_y$, $Cu(OH)_x(CO_3)_y$
Al	γ-Al_2O_3, γ-AlOOH, $Al(OH)_3$ $Al(OH)_xCl_y$, $Al(OH)_x(SO_4)_y$

The formation of basic salts results from the reaction of zinc hydroxide with certain constituents of the atmosphere. For example, the reaction of zinc hydroxide with carbon dioxide leads to formation of weakly soluble hydroxy-carbonates (or basic carbonates) with the general formula $Zn(OH)_x(CO_3)_y$:

$$Zn(OH)_2 + y\, CO_2 \rightarrow Zn(OH)_{2-2y}(CO_3)_y + y\, H_2O \tag{4.48}$$

The value of the coefficient y depends on reaction conditions and does not normally correspond to a whole number. The formula $Zn_5(OH)_6(CO_3)_2$, for example, designates a compound of stoichiometry with $y = 0.4$. The different basic carbonates form a whitish patina and the corrosion rate stays relatively low.

In presence of SO_2 hydroxy-sulfates (basic sulfates) described by the general formula $Zn(OH)_x(SO_4)_y$ can form. These products are weakly soluble in neutral environments, but exhibit a high solubility under acid conditions. In a highly polluted atmosphere, the pH of the electrolyte on the zinc surface can drop to sufficiently low values for the hydroxy-sulfates and the hydroxy-carbonates to dissolve. Rain water running off the surface can carry them away and thus accelerate corrosion. Several studies have shown a linear relationship between the rate of zinc corrosion and the SO_2 concentration in the atmosphere.

In marine atmospheres, the reaction of $Zn(OH)_2$ with chlorides results in hydroxy-chlorides of type $Zn(OH)_xCl_{y'}$ where $y' = 2 - x$. The rate of corrosion of zinc is usually lower than in strongly polluted industrial atmospheres.

Atmospheric corrosion of aluminum

Aluminum corrodes little when exposed to the atmosphere, because it is protected by a thin passive oxide film. The inner part of the film, adjacent to the metal, contains

© 2007, first edition, EPFL Press

mostly alumina (γ-Al_2O_3) while in the outside part one finds boehmite (γ-AlOOH) and bayerite ($Al(OH)_3$).

The presence of chlorides or sulfates leads to the formation of the hydroxy-chlorides $Al(OH)_2Cl$ and $Al(OH)Cl_2$ or of hydroxy-sulfates $Al(OH)_x(SO_4)_y$. These compounds dissolve in acidic environments and the resulting corrosion products are carried away by run-off water. However, even in a polluted atmosphere, the uniform corrosion rate of aluminum remains generally small. On the other hand, in presence of chloride pollution, damage by pitting may occur.

While pits formed on aluminum immersed in a chloride solution can reach an appreciable size and depth (Chapter 7), those formed under atmospheric corrosion remain generally shallow, because periodic drying favors repassivation. According to Table 8.24, the maximum depth of the pits on aluminum alloys exposed to a marine atmosphere for a period of 20 years, did not exceed 260 μm. Although this attack has a negative impact on the visual aspect of a structure, it rarely impedes its function.

Atmospheric corrosion of copper

Upon contact with the atmosphere, copper forms a brown-black patina that becomes blue-green over the years. The changing color is explained as follows. Initially the oxides Cu_2O and CuO are formed. Then, in the presence of pollutants, the oxides slowly transform into greenish hydroxy-sulfates $Cu(OH)_x(SO_4)_y$, hydroxy-carbonates $Cu(OH)_x(CO_3)_y$ or hydroxy-chlorides $Cu(OH)_xCl_y$. Normally the patina has little or no influence on the rate of corrosion, which remains constant during the time of exposure. Uncoated copper structures exposed to the atmosphere, such as for roofing, corrode only very slowly and last many years. More critical than the loss of copper due to atmospheric corrosion is a possible environmental pollution by copper ions formed by corrosion reactions and carried away by rainwater.

Green patina, whose decorative aspect is highly appreciated, contains mostly basic salts. Their exact composition depends on the atmospheric conditions. In urban atmospheres they are mostly composed of hydroxy-sulfates and in marine atmospheres of hydroxy-chlorides. The patina forms more rapidly in polluted atmospheres, but even so, one generally has to wait a few years for a green patina to appear. Environments with a high SO_2 concentration can prevent entirely the formation of the green patina, because the surface pH of the metal remains too low for the precipitation of basic salts.

General Bibliography

F. P. Fehlner, *Low Temperature Oxidation*, Wiley-Interscience New York (1986) 257 pp.

V. Kucera, E. Mattson , *Corrosion Mechanisms*, F. Mansfeld, ed., Marcel Dekker, Inc., New York (1987) pp. 211-284.

C. Leygraf, T. Graedel, *Atmospheric Corrosion*, J. Wiley, NewYork, (2000) 368 pp.

© 2007, first edition, EPFL Press

References

[1] F. P. Fehlner, N. F. Mott, *Oxidation of Metals* 2, 59 (1970).
[2] C. Palacio, H. J. Mathieu, D. Landolt, *Surf. Sci.* 182, 41 (1987).
[3] P. Holloway, *J. Vac. Sci. Technol.* 18, 653 (1981)
[4] A. T. Fromhold, *Theory of Metal Oxydation*, North Holland, Amsterdam (1976) p. 290.
[5] N. F. Mott, *Trans. Farad. Soc.* 39, 472 (1940).
[6] N. Cabrera, N. F. Mott, *Rep. Progr. Phys.* 12, 163 (1948/49).
[7] R. Ghez, *J. Chem. Phys.* 58, 1838 (1973).
[8] F. P. Fehlner, *Low Temperature Oxidation*, Wiley-Interscience, New York (1986) pp. 68-7.
[9] P. Kofstad, *High Temperature Corrosion*, Elsevier, London (1988) p.151.
[10] Norm ISO N. 9223, 1989.
[11] V. Kucera, E. Mattson, *Corrosion Mechanisms*, F. Mansfeld, ed., Marcel Dekker, Inc., New York (1987) pp. 215, 228.
[12] W. H. Vernon, *Trans. Farad. Soc.* 31, 1668 (1935).
[13] U. R. Evans, C. A. J. Taylor, *Corrosion Science* 12, 227 (1972).
[14] K. Barton, *Schutz gegen atmosphärische Korrosion*, Verlag Chemie, Weinheim (1973) p. 73.
[15] H. Schwarz, *Werkstoffe und Korrosion* 23, 648 (1972).
[16] M. Stratmann, *Ber. Bunsenges. Phys. Chem.* 94, 626 (1990).
[17] M. E. Komp, *Materials Performance* 26, 42 (1987).
[18] V. Kucera, E. Mattson, *Corrosion Mechanisms*, F. Mansfeld. ed., Marcel Dekker, Inc., New York (1987) pp. 259, 268, 273.
[19] C. J. Slunder, W. K. Boyd, *Zinc: Its Corrosion Resistance*, Zinc Development Association, London (1971) p. 41.

© 2007, first edition, EPFL Press

CHAPTER 9

HIGH-TEMPERATURE CORROSION

9.1 BASIC CONCEPTS

9.1.1 High-temperature versus low-temperature corrosion

High-temperature corrosion takes place in the absence of aqueous electrolytes, at temperatures at which the rate of solid state diffusion becomes important, typically above 300 °C. High-temperature corrosion is an important problem in many technological activities such as energy production (gas turbines, combustion chambers), transport (jet engines, diesel motors), chemical and metallurgical processing (reactors, furnaces) and waste incineration. In the literature, the term "dry corrosion" is sometimes used to designate high-temperature corrosion, in contrast to "wet corrosion", which refers to corrosion in the presence of an aqueous electrolyte. Since low temperature oxidation (Chapter 8) is also a "dry" process, we will not use this term.

The kinetics of high-temperature corrosion differs from those of corrosion at ambient temperature in three respects:

- no aqueous electrolyte;
- importance of solid state diffusion;
- equilibrium conditions at the interfaces.

Although there are no aqueous electrolytes, high temperature corrosion is an electrochemical process, involving anodic and cathodic partial reactions. The metal oxides generated at the corroding surface or molten salts present at the surface form the electrolyte.

During low temperature oxidation, oxide films grow by high-field conduction (Section 8.1) rather than by solid-state diffusion because the value of the diffusion coefficients is too small. Under these conditions the thickness of the oxide layer does not exceed a few nanometers. In contrast, in high temperature corrosion, volume diffusion and grain-boundary diffusion are the principle transport mechanisms by which oxide layers grow. As a consequence their thickness can reach much larger values.

At ambient temperature, the rate of charge-transfer at the metal-electrolyte interface often limits the corrosion rate (Chapter 4). Because transfer reactions generally exhibit higher activation energy than diffusion phenomena, their rate

© 2007, first edition, EPFL Press

increases more steeply with temperature. As a result, at sufficiently high temperatures diffusion becomes rate controlling. The charge-transfer reactions at the interfaces can then be viewed as being at thermodynamic equilibrium and the overall rate of corrosion is controlled by transport processes in the scale or in the adjacent metal.

9.1.2 Corrosion reactions

At high temperatures, certain gases that are normally considered harmless can undergo a chemical reaction with metals or with non-metallic phases present in an alloy such as carbides. The most common gases found in high temperature corrosion are:

- oxygen;
- water vapor;
- carbon dioxide;
- sulfur dioxide;
- chlorine;
- hydrogen.

Reaction with oxygen

The reaction of a metal with oxygen at temperatures exceeding 300 to 400 °C is referred to as ***high-temperature oxidation***. The properties of the oxide formed at the surface determine the reaction rate.

We can write the reaction stoichiometry in a number of ways. By referring to the formation of one mole of oxide, we obtain an expression that is useful for defining thermodynamic properties, such as the enthalpy of formation.

$$\nu_M \mathrm{M} + \frac{\nu_\mathrm{O}}{2}\,\mathrm{O}_2 = \mathrm{M}_{\nu_\mathrm{M}}\mathrm{O}_{\nu_\mathrm{O}} \tag{9.1}$$

Here ν_M and ν_O represent the stoichiometric coefficients of the metal and of oxygen in the oxide. By referring to the reaction of one mole of metal we obtain equation (9.2). This formulation is commonly used to describe corrosion phenomena.

$$\mathrm{M} + \frac{\nu'}{2}\,\mathrm{O}_2 = \mathrm{M}\,\mathrm{O}_{\nu'} \tag{9.2}$$

The stoichiometric coefficient $\nu' = \nu_\mathrm{O}/\nu_\mathrm{M}$ does not always correspond to a whole number. For example, for the reaction $2\ \mathrm{Al} + 3/2\ \mathrm{O}_2 = \mathrm{Al}_2\mathrm{O}_3$, we find $\nu' = 3/2 = 1.5$.

Expression (9.3), refers to the reaction of one mole of gaseous oxygen. It is obtained by dividing (9.1) by $\nu_\mathrm{O}/2$.

$$2\,\frac{\nu_\mathrm{M}}{\nu_\mathrm{O}}\,\mathrm{M} + \mathrm{O}_2 = \frac{2}{\nu_\mathrm{O}}\,\mathrm{M}_{\nu_\mathrm{M}}\mathrm{O}_{\nu_\mathrm{O}} \tag{9.3}$$

This formulation is particularly useful for the comparison of the thermodynamic stability of the oxides of different metals using *Ellingham diagrams* (Figure 2.2).

© 2007, first edition, EPFL Press

Reaction with water vapor

At high temperature, a reaction equilibrium between a metal and water vapor may be established as follows:

$$M + \nu' H_2O = MO_{\nu'} + \nu' H_2 \tag{9.4}$$

With the data of Figure 2.2 or of tables of free energies of formation, the equilibrium partial pressures can be calculated.

The *Schikorr reaction* (9.5) describes corrosion of iron in contact with water vapor at temperatures up to 600 °C.

$$3\,Fe + 4\,H_2O = Fe_3O_4 + 4\,H_2 \tag{9.5}$$

The Schikorr reaction results in the formation of a compact layer of magnetite whose growth governs the kinetics of corrosion. Below 300 °C, the rate of film growth is very slow and the magnetite film effectively protects against further corrosion (Section 8.1).

Reaction with carbon dioxide

The carbon dioxide can oxidize a metal according to the following reaction equilibrium:

$$M + \nu' CO_2 = MO_{\nu'} + \nu' CO \tag{9.6}$$

According to (9.6) the corrosivity of the gas depends on the ratio P_{CO2}/P_{CO}. Exhaust gas from combustion processes generally contains little CO, thus favoring the reaction of the metal with carbon dioxide. In contrast, in a blast furnace containing iron oxide and coke, the partial pressure of CO is elevated. Under these conditions, iron oxide is reduced into iron. This shows that the same reaction can cause corrosion of a metal or serve for its elaboration.

Reaction with sulfur compounds

Fossil fuels (oil, coal) contain organic sulfur compounds that, during combustion, are transformed into SO_2 and SO_3. At high temperature, the reaction of these gases with metals yields sulfides, oxides or sulfates:

$$3\,M + \nu' SO_2 = MS_{\nu'} + 2\,MO_{\nu'} \tag{9.7}$$

$$M + \nu' SO_3 = MO_{\nu'} + \nu' SO_2 \tag{9.8}$$

$$M + \nu' SO_2 + \nu' O_2 = M(SO_4)_{\nu'} \tag{9.9}$$

Under reducing conditions, metals can react with gaseous sulfur and hydrogen sulfide (H_2S) according to:

© 2007, first edition, EPFL Press

$$M + \frac{\nu'}{2} S_2 = M S_{\nu'} \tag{9.10}$$

$$M + \nu' H_2S = MS_{\nu'} + \nu' H_2 \tag{9.11}$$

Because of their low melting points, much lower that those of the oxides, sulfides and sulfates are often present in the liquid state on metal surfaces. They destabilize the oxide films and thus accelerate the rate of corrosion.

Reaction with chlorine

Gaseous chlorine is a powerful oxidant that reacts with metals according to the following equation:

$$M + \frac{\nu'}{2} Cl_2 = M Cl_{\nu'} \tag{9.12}$$

The practical importance of this reaction remains limited. On the other hand, exhaust gases from combustion processes may contain chlorides that originate from the fuel or from the intake air. Well-known examples are incineration of domestic waste that contains chlorides and intake of maritime air by jet engines. Chlorides carried by exhaust gases can deposit on metal surfaces and together with other salts form highly corrosive ionic liquids.

Reaction with hydrogen

Under certain conditions, hydrogen reacts with carbides present in a metal to form methane. The ***decarburization of steel*** is a reaction of hydrogen with iron carbide Fe_3C (cementite):

$$Fe_3C + 4\,H_m = 3\,Fe + CH_4 \tag{9.13}$$

In this equation, H_m refers to atomic hydrogen dissolved in the metal. Reaction (9.13) has been identified as causing damage to installations exposed to pressurized hydrogen at elevated temperatures in the chemical and petroleum industry. The reaction weakens the mechanical strength of the metal and represents a risk of accident.

Certain alloying elements of steel form carbides that are more stable than iron carbide. Their presence helps to prevent the decarburization. Figure 9.1 gives the standard free energy of formation for different carbides, calculated for the reaction of one mole of carbon in the form of graphite. The values for chromium and molybdenum carbides are more negative than those for iron carbide, indicating a greater thermodynamic stability. This explains why chromium- and molybdenum-containing steels are more resistant to decarburization than carbon steels. Figure 9.2, known as *Nelson diagram*, indicates the limiting temperature for use of different steels in the presence of hydrogen.

© 2007, first edition, EPFL Press

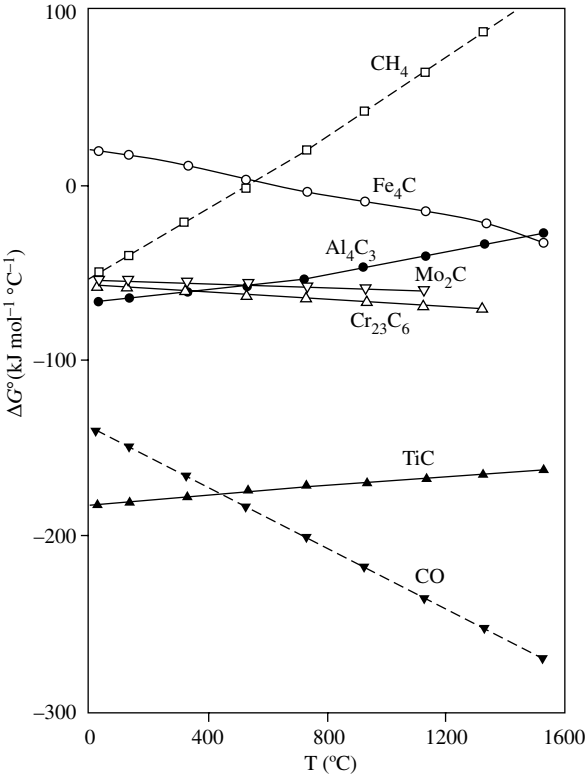


Figure 9.1 Variation of the standard free energies of formation for carbides as a function of temperature. The values listed refer to the reaction written for one mole of carbon (Ellingham diagram). This diagram was prepared using the data found in ref. [1].

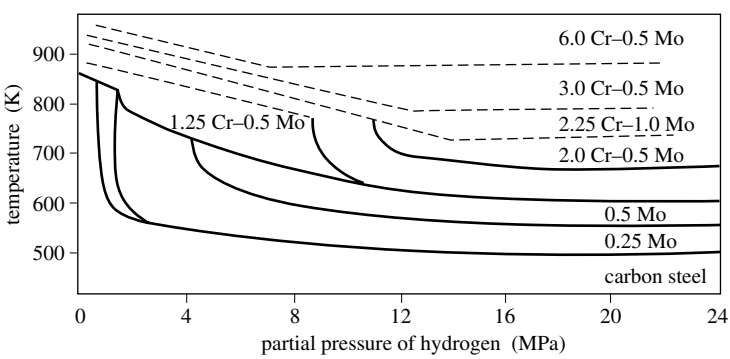


Figure 9.2 Nelson diagram showing the limits for use of different Cr-Mo steels in the chemical process industry[2].

© 2007, first edition, EPFL Press

Under certain conditions, dissolved oxygen (O_m) can cause the decarburization of steel by oxidizing iron carbide into CO.

$$Fe_3C + O_m = 3\,Fe + CO \tag{9.14}$$

Because in the metal oxygen diffuses less readily than hydrogen, this reaction normally takes place only near the surface. During steel production, it can lead to carbon depletion at the surface, requiring a thermal treatment to restore a uniform carbide concentration. From a corrosion point of view the reaction (9.14) is usually not critical.

9.1.3 Experimental study of high-temperature corrosion

Thermogravimetry is the most commonly used method for studying the kinetics of corrosion reactions at high temperatures. The sample is placed in an oven with a controlled atmosphere, where one monitors the variation of its mass by means of a microbalance (Figure 9.3). The precision of the technique is on the order of micrograms. If all corrosion products remain on the surface of the sample, the measured mass increase as a function of time permits to identify the rate law that governs the reaction kinetics.

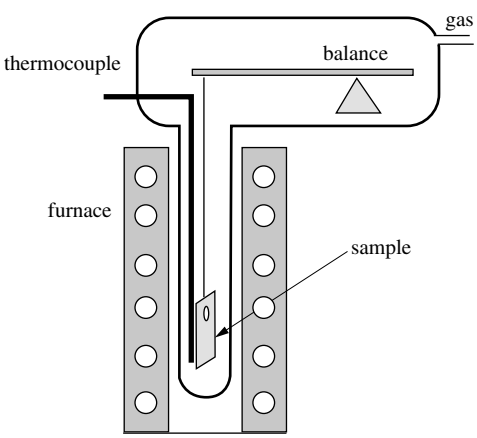


Figure 9.3 Schematic showing the measurement of corrosion rate by thermogravimetry.

Oxidation, the chemical reaction of a metal with oxygen, is the most studied reaction in high-temperature corrosion. It shows two limiting behaviors that can be described by simple models (Figure 9.4):

- the linear growth law;
- the parabolic growth law.

© 2007, first edition, EPFL Press

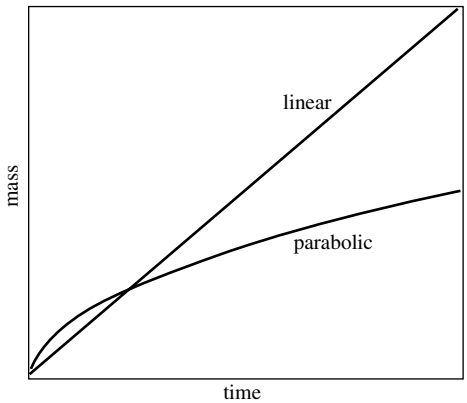


Figure 9.4 Linear and parabolic rate laws of oxidation.

Linear growth law

When the rate of oxidation is constant,

$$\frac{\mathrm{d}m_{\mathrm{oxd}}}{\mathrm{d}t} = k'_{\mathrm{p}} \tag{9.15}$$

the mass of the oxide formed increases linearly with oxidation time:

$$m_{\mathrm{oxd}} = k'_{\mathrm{p}}\, t \tag{9.16}$$

In these equations, m_{oxd} designates the mass of the oxide formed per unit of surface area and k'_{p} is the ***linear oxidation constant*** (kg m^{-2} s^{-1}). Equation (9.15), which follows from integration of (9.15) with the initial condition $m_{\mathrm{oxd}} = 0$ at $t = 0$, is referred to as the ***linear growth law***. A linear mass increase indicates that the accumulation of oxidation products at the surface has no effect on the reaction rate, in other words, the oxide layers formed are porous or cracked.

Parabolic growth law

When the transport of ionic species across the oxide layer limits the rate of reaction, the oxidation rate is inversely proportional to the mass of oxide forming on the surface per unit area:

$$\frac{\mathrm{d}m_{\mathrm{oxd}}}{\mathrm{d}t} = \frac{k_{\mathrm{p}}}{m_{\mathrm{oxd}}} \tag{9.17}$$

Here k_{p} is the ***parabolic oxidation constant*** (kg^2 m^{-4} s^{-1}). Integration of (9.17) with the condition $m_{\mathrm{oxd}} = 0$ at $t = 0$, yields the ***parabolic growth law***:

$$m_{\mathrm{oxd}} = \left(2\, k_{\mathrm{p}}\right)^{1/2} t^{1/2} \tag{9.18}$$

© 2007, first edition, EPFL Press

In reality, the functions $m_{oxd} = f(t)$ are often more complex than described by the two limiting models. We have seen in Chapter 8 that low-temperature oxidation generally follows a logarithmic or an inverse logarithmic rate law. At intermediate temperatures, the observed growth laws may be between parabolic and logarithmic behavior. Finally, during high temperature oxidation, the structure and the composition of the oxide layers can change during the course of a reaction, for example due to the presence of sulfur. The reaction kinetics then no longer follow a simple growth law and the oxidation curves exhibit discontinuities due to loss of adhesion or to cracking of the oxide film.

To get mechanistic information, gravimetric measurements should always go together with metallographic and analytical investigations aimed at the study of the morphology and the composition of the corrosion products and the metal in the vicinity of the surface. For this, a scanning electron microscope equipped with an X-ray microanalysis system is particularly well suited. Transmission electron microscopy in conjunction with electron diffraction, and conventional X-ray diffraction methods are also frequently applied. Surface analysis methods such as XPS, AES and SIMS permit the study of the oxide-metal interface and of initial oxide growth.

9.2 OXIDATION OF METALS

9.2.1 Structure and properties of oxides

The structure and the physical properties of oxides determine the oxidation kinetics at high temperature. Table 9.5 provides a selection of relevant data.

Crystal structure

Common metal oxides most often exhibit a cystal structure that corresponds to a compact arrangement of oxygen anions. The metal cations are located at the tetrahedral and octahedral sites of the anion lattice. On average, there are two tetrahedral and one octahedral site per oxygen ion.

In ***divalent oxides*** of type MO (e.g., MgO, CoO, NiO, FeO, TiO, NbO, VO), the cations generally occupy the octahedral sites. These oxides crystallize in the NaCl structure, although there are some exceptions to this rule. The oxides ZnO and BeO, for example, have the wurtzite structure in which the cations occupy half of the tetrahedral sites.

A large number of ***trivalent oxides*** with the general formula M_2O_3 have the structure of corundum α-Al_2O_3. Examples include α-Fe_2O_3, Cr_2O_3, Ti_2O_3 and V_2O_3. The metal cations in these oxides occupy two thirds of the octahedral sites.

The ***spinels*** are oxides with the general formula $M^{(II)}M_2^{(III)}O_4$ in which the sum of the cationic valences is +8. Examples include $MgAl_2O_4$ ("spinel"), $FeCr_2O_4$ and $CoCr_2O_4$. In spinels, the cations occupy the tetrahedral and octahedral sites of the oxygen ion cubic lattice. In $MgAl_2O_4$ the divalent Mg atoms occupy the tetrahedral sites and the trivalent Al atoms occupy the octahedral sites. ***Magnetite***, with the stoichiometric formula Fe_3O_4, contains both divalent and trivalent iron cations. The

© 2007, first edition, EPFL Press

Table 9.5 Physical properties of some oxides [3].

Oxide	Crystal structure	Melting point (°C)	Boiling or decomposition temperature (°C)	Molar volume ($cm^3\ mol^{-1}$)	R_{PB}	Thermal expansion coefficient $\times 10^6$	R_{CD}
Li_2O	CaF_2	1700	3092	14.8	0.57		
Cu_2O	cubic	1235	2255	23.8	1.67	4.3	0.23
CoO	NaCl	1935	3515	11.6	1.74	15.0	1.07
FeO	NaCl	1420	2588	12.6	1.78	12.2	0.80
NiO	NaCl	1990	3614	11.2	1.70	17.1	0.97
MgO	NaCl	2800	5072	11.3	0.8		
α-Al_2O_3	corundum	2015	3659	25.7	1.28		
Cr_2O_3	corundum	2535	4415	29.2	2.02	7.3	0.77
α-Fe_2O_3	corundum	1565	2849	30.5	2.15	14.9	0.97
γ-Fe_2O_3	cubic	1457	2655	31.5	2.22		
Fe_3O_4	spinel			44.7	2.10		
TiO_2	rutile	1830	3326	18.8	1.76		
ZrO_2	monocl.	2715	4919	22.0	1.57		
Ta_2O_5	tricl.	1800	3272	53.9	2.47		
V_2O_5	orthorhomb	690	1274	54.2	3.25		
MoO_3	orthorhomb.	795	1463	30.7	3.27		
WO_3	orthorhomb.	1473		32.4	3.39		

divalent ions occupy the octahedral sites and the trivalent cations part of the tetrahedral sites. Such a structure is called *inverted spinel* structure ($FeFe_2O_4$).

The tetravalent oxides with the general formula MO_2 often have the ***rutile*** structure. Besides TiO_2 ("rutile"), they include SnO_2, MnO_2 and MoO_2. In this structure, the metal ions occupy the center of the octahedron formed of oxygen anions. The anions are arranged in such a way that each of them belongs to three octahedra. The structures of higher valent oxides, with the general formula of M_2O_5 or MO_3, are in general more complicated.

Melting temperature

The oxides of metals or alloys with good resistance to high temperature corrosion generally have a high melting temperature. They therefore remain in the solid state under typical working conditions. However, there are several exceptions to this rule.

According to Table 9.5, MoO_3 and V_2O_5 melt at 795 °C and 690 °C, respectively. Once liquid, these oxides can dissolve other oxides that normally resist at these temperatures, and thus they accelerate the corrosion rate. The accelerated corrosion of gas turbines due to V_2O_5 present in combustion gases illustrates this effect. The V_2O_5 which has a boiling point of 1770 °C vaporizes during combustion and is

© 2007, first edition, EPFL Press

present in the hot gas that drives the turbine. It condenses on the metal surfaces whose temperature is lower, and causes fluxing of the oxide scales formed.

Chromium is an important component of the so-called superalloys that resist well to high-temperature corrosion. The oxide Cr_2O_3 exhibits a high melting temperature of 2535 °C. At high partial pressures of oxygen, typically above 0.2 bar, and elevated temperatures, the trivalent oxide Cr_2O_3 may oxidize into the hexavalent oxide CrO_3. The CrO_3 is quite volatile and evaporates above 1000 °C. As a consequence, the alloys lose their resistance to corrosion.

Molar volume

Because the molar volumes of oxides differ from those of the corresponding metals, the oxide layers formed are subject to tensile or compressive stress. If sufficiently important, these can lead to cracking or pealing of the oxide films. The ***Pilling-Bedworth Ratio,*** $R_{PB,}$ is a measure of the relative magnitude of the molar volumes of oxides and metals.

$$R_{PB} = \frac{V_{oxd}}{\nu_M V_M} \tag{9.19}$$

Here, V_{oxd} and V_M represent the molar volumes of the oxide and of the metal, respectively, and ν_M designates the stoichiometric coefficient of the metal for the formation of one mole of oxide (equation (9.1)).

In principle, if $R_{PB} \approx 1$, the stresses remain weak, and the oxide film should therefore well protect the metal against high-temperature corrosion. However, practical experience does not always agree with this prediction. The reason is that the Pilling-Bedworth ratio is just one of several criteria that determine the stress in oxide films. Among other, it does not allow for the capacity of the metal-oxide system to accommodate stresses by elastic or plastic deformation; nor does it take into account the influence of the growth mechanism. Indeed, a metal oxide system can in general more easily tolerate a difference in molar volume if the oxide grows by cation diffusion from the metal towards the outside surface, rather than by anion diffusion in the opposite direction.

Thermal expansion coefficient

The ratio R_{CD} of the thermal expansion coefficients of the oxide, α_{oxd}, and of the metal, α_m, determines the magnitude of the stresses that develop during thermal cycling.

$$R_{CD} = \frac{\alpha_{oxd}}{\alpha_m} \tag{9.20}$$

The thermal expansion coefficient of oxides is generally smaller that that of the corresponding metals, because of their higher melting temperature. Cooling therefore

© 2007, first edition, EPFL Press

induces compressive stress that can cause the oxide film to peal away from the substrate. The importance of this effect depends on a number of parameters, including the geometry of the sample.

Concentration of point defects in stoichiometric oxides

Diffusion in oxides occurs by means of point defects: interstitial ions and vacancies. In ***stoichiometric oxides*** one finds two types of point defects, referred to as Frenkel defects and Schottky defects. These are presented schematically, for a divalent oxide of type MO, in Figure 9.6. To describe the chemical reactions of point defects and electronic charges we use the symbols given in Table 9.7. A ***Frenkel defect*** in a divalent oxide consists of an interstitial cation $M_i^{\bullet\bullet}$ and a cation vacancy V_M''. A ***Schottky defect*** consists of a cation vacancy V_M'' and an anion vacancy $V_O^{\bullet\bullet}$. Because the oxide as a whole is electrically neutral, the concentrations of negative and positive charges associated with the defects are equal (electroneutrality condition). Thus, we find for Frenkel defects,

$$c_{M_i^{\bullet\bullet}} = c_{V_M''} \tag{9.21}$$

and for Schottky defects,

$$c_{V_M''} = c_{V_O^{\bullet\bullet}} \tag{9.22}$$

In stoichiometric oxides, the equilibrium concentration of the defects depends only on temperature. For Frenkel defects at equilibrium:

$$c_{M_i^{\bullet\bullet}}\, c_{V_M''} = c^2_{M_i^{\bullet\bullet}} = c^2_{V_M''} = K \tag{9.23}$$

In this equation, the equilibrium constant K is equal to: $K = \exp(-\Delta G_f / RT)$, where ΔG_f is the free energy of formation of the defects. Analogous expressions can be written for Schottky defects and for electronic charge carriers (electrons and holes).

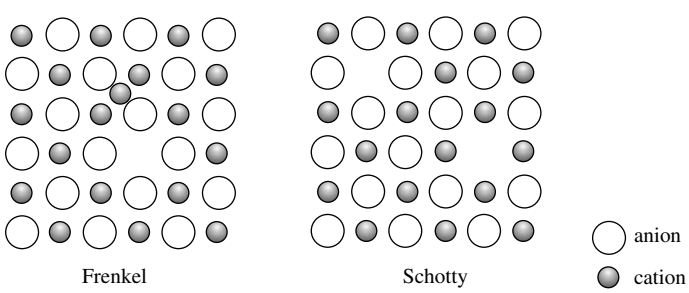


Figure 9.6 Point defects in MO-type oxides.

© 2007, first edition, EPFL Press

Table 9.7 Designation of point defects in oxides.

Defect	One positive charge	Two positive charges	One negative charge	Two negative charges
Cation vacancy			V'_M	V''_M
Anion vacancy	$V^{\bullet}_O$	$V^{\bullet\bullet}_O$		
Interstitial cation	$M^{\bullet}_i$	$M^{\bullet\bullet}_i$		
Interstitial anion			O'_i	O''_i
Electronic charge	h (subscript: p)		e (subscript: n)	

Concentration of point defects in non-stoichiometric oxides

Non-stoichiometric oxides have a slight excess of either metal cations or oxygen anions. Their general formula corresponds to $M_{1+x}O_{v'}$ and $M_{1-x}O_{v'}$, respectively, where x is a positive number such that x <<1. The charge associated with the excess ions is compensated for by free electrons or holes.

Thus, in the oxides $M_{1+x}O_{v'}$, electrons compensate the positive charge resulting from the excess of cations. These oxides behave as *n*-type semiconductors. Examples are: ZnO, CdO, MgO TiO_2, V_2O_5, Fe_2O_3, Ta_2O_5, MoO_3 , SnO_2.

In the oxides $M_{1-x}O_{v'}$, electron holes compensate the negative charge due to the excess of anions. These oxides behave as *p*-type semiconductors. Some examples: Cr_2O_3, $FeCr_2O_4$, MnO, FeO, NiO, CoO, Cu_2O. (Figure 9.8)

In non-stoichiometric oxides, the concentration of point defects at equilibrium depends on the partial pressure of oxygen. For a divalent ***n-type oxide*** with interstitial cations $M^{\bullet\bullet}_i$, the equilibrium with oxygen is expressed by the stoichiometric equation:

$$MO = \frac{1}{2}O_2 + M^{\bullet\bullet}_i + 2\,e \tag{9.24}$$

Figure 9.8 schematically shows this reaction. The electroneutrality condition gives:

$$2\,c_{M^{\bullet\bullet}_i} = c_n \tag{9.25}$$

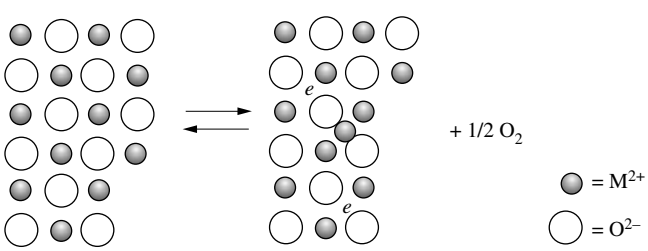


Figure 9.8 Equilibrium between a bivalent *n*-type oxide and oxygen according to (9.24).

© 2007, first edition, EPFL Press

Here, $c_{M_i^{\bullet\bullet}}$ and c_n designate, respectively, the concentrations of interstitial cations and of electrons. Given the low concentration of point defects, we can replace activities by concentrations. The equilibrium constant then becomes:

$$K = c_{M_i^{\bullet\bullet}} \; c_n^2 \; P_{O_2}^{1/2} = 4\, c_{M_i^{\bullet\bullet}}^3 \; P_{O_2}^{1/2} \tag{9.26}$$

and the equilibrium concentration of interstitial cations is equal to:

$$c_{M_i^{\bullet\bullet}} = \left(\frac{K}{4}\right)^{1/3} P_{O_2}^{-1/6} \tag{9.27}$$

In an analogous fashion, for a divalent ***p-type oxide*** the equilibrium between oxygen and cation vacancies V_M'' is described by the stoichiometric equation:

$$\frac{1}{2}\, O_2 = MO + V_M'' + 2\, h \tag{9.28}$$

This equilibrium is schematically presented in Figure 9.9. With the electroneutrality condition,

$$2\, c_{V_M''} = c_p \tag{9.29}$$

where c_p indicates the concentration of the holes, we obtain:

$$K = c_{V_M''} \; c_p^2 \; P_{O_2}^{-1/2} = 4\, c_{V_M''}^3 \; P_{O_2}^{-1/2} \tag{9.30}$$

and

$$c_{V_M''} = \left(\frac{K}{4}\right)^{1/3} P_{O_2}^{1/6} \tag{9.31}$$

It follows that with increasing oxygen partial pressure the concentration of ionic vacancies increases in *p*-type oxides, whereas it decreases in *n*-type oxides. By applying analogous reasoning to the case of monovalent defects in an oxide of type M_2O, it is easily shown that:

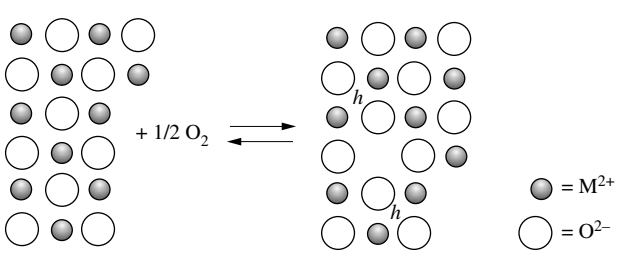


Figure 9.9 Equilibrium between a bivalent *p*-type oxide and oxygen according to (9.28).

© 2007, first edition, EPFL Press

$$c_{M_i^{\bullet}} \propto P_{O_2}^{-1/8} \tag{9.32}$$

and

$$c_{V_M'} \propto P_{O_2}^{1/8} \tag{9.33}$$

Concentration of point defects in doped oxides

In doped semiconductors, the concentration of charge carriers is given by the concentration of the dopant, which may be an electron acceptor or an electron donor. Cations that have a valence lower than that of the base oxide are electron acceptors. In a divalent oxide MO, monovalent cations such as Li^+ or K^+ act as electron acceptors.

Indeed, if we replace an M^{2+} cation in the MO lattice by Li^+ or some other monovalent cation, a positive charge will be missing. We therefore assign a virtual negative charge to the defect. To compensate this negative charge, the metal cations located at interstitial positions provide a positive charge. If $c_{A'}$ indicates the acceptor concentration, the electroneutrality condition yields:

$$c_{M_i^{\bullet\bullet}} = \frac{1}{2} c_{A'} \tag{9.34}$$

In contrast, ***electron donors*** have a higher valence than that of the base oxide. In a divalent oxide of type MO, trivalent cations such as Al^{3+} or $Cr^{3+}$are electron donors. When, in such an oxide, an Al^{3+} ion replaces M^{2+}, a virtual positive charge appears. This charge, assigned to the defect, is compensated for by the formation of cation vacancies. If $c_{D^{\bullet}}$ designates the donor concentration, the concentration of cation vacancies is equal to:

$$c_{V_M''} = 1/2\, c_{D^{\bullet}} \tag{9.35}$$

In summary, the concentration of point defects in stoichiometric oxides depends only on the temperature, whereas for non-stoichiometric oxides, it also depends on the partial pressure of oxygen. Finally, in doped oxides, it is determined by the dopant concentration.

9.2.2 Wagner's theory of oxidation

The growth of compact oxide films, when it obeys the parabolic rate law, implies that M^{n+} cations or O^{2-} anions are moving across the oxide (Figure 9.10). The oxides of most metals or alloys exhibiting good resistance to high-temperature corrosion (nickel, cobalt, iron, chromium, etc.), grow by cation diffusion. Only this mechanism will be discussed in the following.

The ***diffusion of cations*** involves movement of point defects: interstitial cations in *n*-type oxides or cation vacancies in *p*-type oxides (Figure 9.11). In order to

© 2007, first edition, EPFL Press

maintain electroneutrality, electrons or holes move simultaneously with the ionic defects. Figure 9.12 provides an overview of defect motion during the growth of a divalent oxide film MO. The cation vacancies advance in the direction opposite to that of cation motion.

Wagner's theory of oxidation provides a quantitative description of the growth rate of compact oxide layers as a function of the difference in electrochemical potential between the metal-oxide and the oxide-gas interfaces. The following analysis uses concepts developed in Section 4.3 for aqueous electrolytes. This simplifies the theoretical developments proposed by Wagner [4], while yielding the same results.

According to equation (4.123), the flux of *i*-type defects in an oxide is proportional to their concentration c_i, to their mobility u_i, and to the electrochemical potential gradient perpendicular to the surface, $d\mu_i^*/dy$:

$$N_i = -c_i u_i \frac{d\mu_i^*}{dy} = -c_i u_i \left(\frac{d\mu_i}{dy} + z_i F \frac{d\Phi}{dy} \right) \tag{9.36}$$

By replacing the activities by concentrations, and using the Nernst-Einstein equation (4.127), we obtain an expression analogous to (4.126):

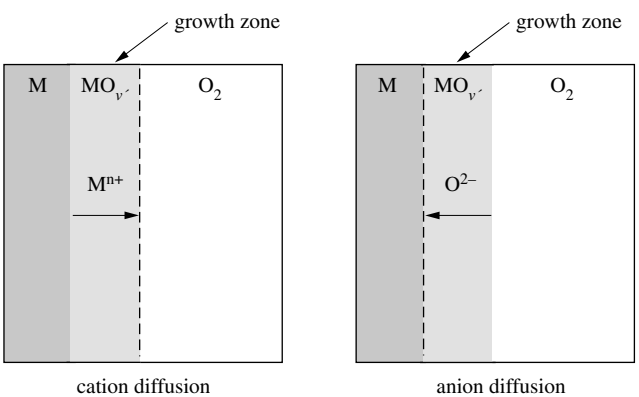


Figure 9.10 Oxide-layer growth by diffusion of cations or anions.

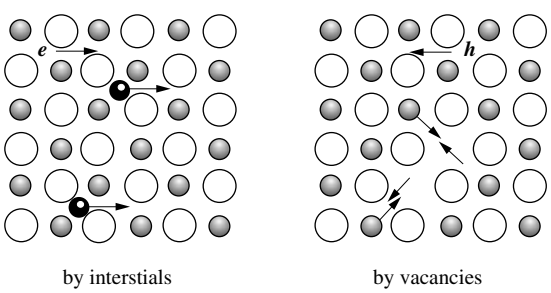


Figure 9.11 Cation diffusion in oxide MO by interstitials or vacancies.

© 2007, first edition, EPFL Press

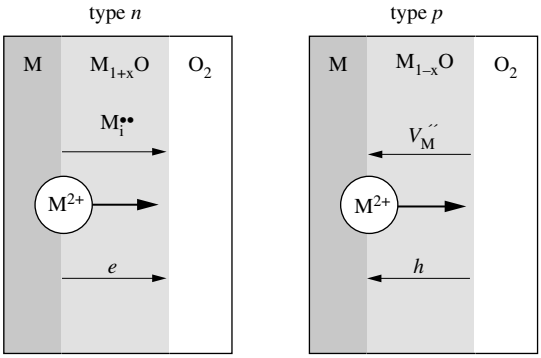


Figure 9.12 Cation diffusion mechanisms during growth of $M_{1\pm x}O$ binary oxides.

$$N_i = -D_i \frac{dc_i}{dy} - z_i F c_i u_i \frac{d\Phi}{dy} \tag{9.37}$$

This equation indicates that the ionic defects move through the oxide by diffusion and migration, just as we found in section 4.3 for ionic transport in aqueous solutions. Equation (9.37) holds for all defect fluxes in the oxide. However, an analytical solution is found only for a single case: the simultaneous movement of two types of mobile defects [6].

In the following, we will develop Wagner's theory of oxidation for divalent oxides with an excess of either cations or anions:

- *n*-type oxides $M_{1+x}O$, that grow by diffusion of interstitial cations and electrons;
- *p*-type oxides $M_{1-x}O$, that grow by diffusion of cation vacancies and holes.

Growth of the n-type oxides $M_{1+x}O$

Equation (9.37), applied to divalent interstitial cations (i =1) and to electrons (i =2), yields:

$$N_1 = -D_1 \frac{dc_1}{dy} - z_1 F c_1 u_1 \frac{d\Phi}{dy} \tag{9.38}$$

$$N_2 = -D_2 \frac{dc_2}{dy} - z_2 F c_2 u_2 \frac{d\Phi}{dy} \tag{9.39}$$

The oxide is electronically neutral everywhere. The fluxes of species 1 and 2 are therefore coupled; in other words, as many positive charges as negative charges must move in order to maintain electroneutrality.

$$z_1 N_1 + z_2 N_2 = 0 \tag{9.40}$$

© 2007, first edition, EPFL Press

The electroneutrality condition states:

$$z_1 c_1 + z_2 c_2 = 0 \tag{9.41}$$

and hence:

$$c_2 = -\frac{z_1}{z_2} c_1 \tag{9.42}$$

Solving (9.39) for $\mathrm{d}\Phi/\mathrm{d}y$, we can eliminate the potential gradient from (9.38). After rearrangement we get:

$$N_1 = \frac{D_1 z_2 u_2 - D_2 z_1 u_1}{z_1 u_1 - z_2 u_2} \frac{\mathrm{d}c_1}{\mathrm{d}y} \tag{9.43}$$

The mobility of electrons far exceeds that of the interstitial cations ($u_2 >> u_1$). The equation (9.43) can therefore be simplified to yield (9.44), where we used the Nernst-Einstein equation (4.127) to replace the mobilities by the diffusion coefficients.

$$N_1 = -\frac{z_2 - z_1}{z_2} D_1 \frac{\mathrm{d}c_1}{\mathrm{d}y} \tag{9.44}$$

According to this equation, the flux of interstitial cations N_1 is that expected for a simple diffusion process, except for a multiplication factor, which in the present case is equal to:

$$\frac{z_2 - z_1}{z_2} = \frac{(-1)-(+2)}{(-1)} = 3 \tag{9.45}$$

The integration of (9.44) over a layer of thickness L yields:

$$N_1 = -3 D_1 \frac{c_{1,\mathrm{g}} - c_{1,\mathrm{m}}}{L} \tag{9.46}$$

In this equation, $c_{1,\mathrm{g}}$ and $c_{1,\mathrm{m}}$ designate the concentrations of interstitial cations at the oxide-gas and the metal-oxide interfaces, respectively.

The layer grows at a rate proportional to the flux N_1 and the molar volume of the oxide V_{oxd}.

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \frac{N_1 V_{\mathrm{oxd}}}{\nu_{\mathrm{M}}} \tag{9.47}$$

The stoichiometric coefficient ν_{M} indicates the number of moles of metal required to form one mole of oxide. For an oxide MO, we have $\nu_{\mathrm{M}} = 1$.

According to the parabolic growth law (9.17), the rate of oxidation is inversely proportional to the mass of the oxide present, m_{oxd}. If the density of the oxide ρ_{oxd} is independent of the layer thickness, we can set:

$$m_{\mathrm{oxd}} = \rho_{\mathrm{oxd}} L A \tag{9.48}$$

© 2007, first edition, EPFL Press

where A represents the exposed surface and L is the thickness of the oxide. With (9.17) and for $A = 1\ \mathrm{m}^2$ we get the following relation:

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \frac{k_\mathrm{p}}{\rho_\mathrm{oxd}^2\, L} \tag{9.49}$$

Combining equations (9.46), (9.47) and (9.49) we obtain an expression for the growth constant of an $M_{1+x}O$ oxide:

$$k_\mathrm{p} = -3\, D_1\, V_\mathrm{oxd}(\mathrm{c}_{1,\mathrm{g}} - \mathrm{c}_{1,\mathrm{m}})\, \rho_\mathrm{oxd}^2 \tag{9.50}$$

In this equation, the subscript 1 indicates the interstitial cation $M_i^{\bullet\bullet}$.

At the oxide-gas interface, the concentration of interstitial cations is normally very low: $c_{1,\mathrm{g}} << \mathrm{c}_{1,\mathrm{m}}$. By setting, $V_\mathrm{oxd}\rho_\mathrm{oxd} = M_\mathrm{oxd}$, where M_oxd represents the molecular mass of the oxide, and by substituting the subscript 1 by $M_i^{\bullet\bullet}$, the expression (9.50) becomes:

$$k_\mathrm{p} = 3\, D_{M_i^{\bullet\bullet}}\ M_\mathrm{oxd}\ \rho_\mathrm{oxd}\ c_{M_i^{\bullet\bullet},\mathrm{m}} \tag{9.51}$$

According to this equation, the parabolic oxidation constant depends on the diffusion coefficient of the interstitial cations and on the concentration of these at the oxide-metal interface.

Growth of p-type oxides $M_{1-x}O$

The equation (9.44) is equally applicable to the growth of the $M_{1\text{-}x}O$ p-type oxides, which is limited by the transport of cation vacancies and holes. The subscript i =1 in this case designates the cation vacancies having a charge of $z_1 = -2$; and the subscript $i = 2$ designates the holes with charge $z_2 = 1$. Because the flux of holes and vacancies must be in opposite directions, equation (9.52) replaces the equation (9.47).

$$\frac{\mathrm{d}L}{\mathrm{d}t} = -\frac{N_1 V_\mathrm{oxd}}{v_\mathrm{M}} \tag{9.52}$$

The parabolic oxidation constant is then equal to:

$$k_\mathrm{p} = 3\, D_1 V_\mathrm{oxd}(c_{1,\mathrm{g}} - c_{1,\mathrm{m}})\, \rho_\mathrm{oxd}^2 \tag{9.53}$$

At the metal-oxide interface, the concentration of cation vacancies is negligible, $c_{1,\mathrm{m}} << \mathrm{c}_{1,\mathrm{g}}$. By setting $V_\mathrm{oxd}\ \rho_\mathrm{oxd} = M_\mathrm{oxd}$, and by substituting the subscript 1 by V_M'' we find:

$$k_\mathrm{p} = 3\, D_{V_\mathrm{M}''}\ M_\mathrm{oxd}\ c_{V_\mathrm{M}'',\mathrm{g}}\ \rho_\mathrm{oxd} \tag{9.54}$$

The parabolic oxidation constant in this case is proportional to the diffusion coefficient of the cation vacancies and to their concentration at the oxide-gas interface.

© 2007, first edition, EPFL Press

Influence of the oxygen partial pressure

By studying how the parabolic oxidation constant varies with the partial pressure of oxygen we can identify the transport mechanism governing the growth of a given oxide film. In equation (9.51), applicable to the $M_{1+x}O$ oxides that grow by movement of interstitial cations, the concentration of the interstitial cations at the metal-oxide interface, $c_{M_{i,m}^{\bullet\bullet}}$, does not vary with the oxygen partial pressure. The parabolic oxidation constant corresponding to this mechanism therefore is independent of the oxygen partial pressure.

In contrast, in equation (9.54), applicable to the $M_{1-x}O$ oxides that grow by movement of cation vacancies, the vacancy concentration at the oxide-gas interface, $c_{V_{M,g}''}$, is proportional to $P_{O2}^{1/6}$. The parabolic oxidation constant, therefore, is also proportional to $P_{O2}^{1/6}$. Figure 9.13 presents the parabolic oxidation constant for nickel as a function of the logarithm of the oxygen partial pressure. For different temperatures, the observed slopes are close to 1/6. This is in agreement with the theoretical prediction for a *p*-type oxide corresponding to the formula $Ni_{1-x}O$.

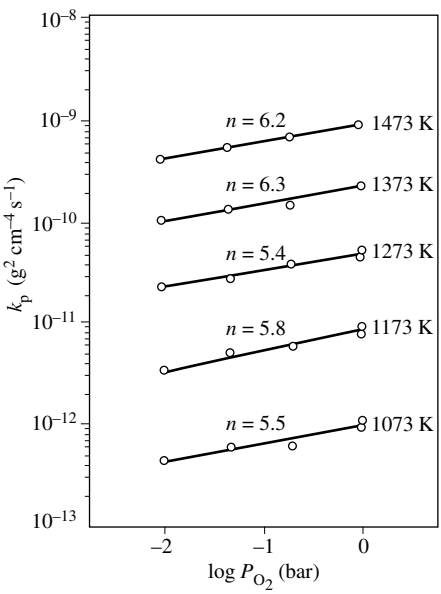


Figure 9.13 Variation of parabolic oxidation constant of Ni with oxygen partial pressure at different temperatures [6].

Influence of the oxide bond energy

The value of the parabolic oxidation constant varies from one metal to another, because it depends on the bond energy and the structure of the oxides formed. In Figure 9.14 the parabolic oxidation constants k_p for different oxides are plotted versus temperature. The values for Al_2O_3 and SiO_2 are particularly small because

© 2007, first edition, EPFL Press

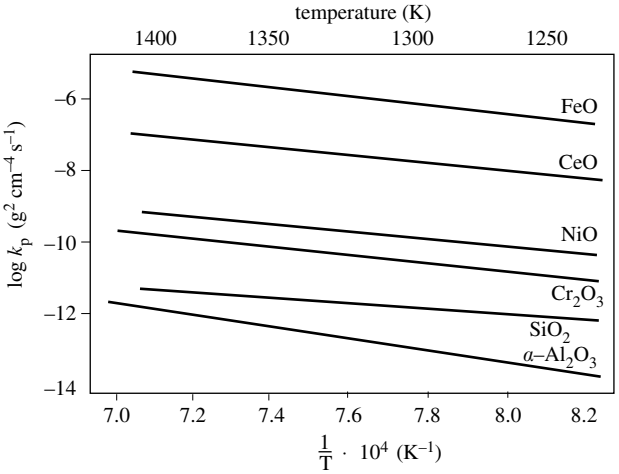


Figure 9.14 Temperature variation of parabolic rate constants for the growth of different oxides [7].

these oxides are thermodynamically very stable. As a consequence, addition of aluminum and silicon to alloys is a means to improve their resistance to high-temperature corrosion (Section 9.3). The values of k_p for the oxides of chromium and nickel, the main components of numerous oxidation-resistant alloys, are also relatively small.

9.2.3 Deviations of Wagner's oxidation theory

The oxidation behavior or metals under operating conditions does not always follow Wagner's theory. Several reasons are responsible for this:

- diffusion at grain boundaries;
- the formation of multiple layers;
- growth stresses.

Diffusion at grain boundaries

Wagner's oxidation theory assumes that *volume diffusion* of point defects limits the growth of oxide layers. However, other transport mechanisms are possible, notably grain boundary diffusion. At relatively low temperatures, $T < 1/2\ T_f$, where T_f is the melting temperature of the oxide, this mechanism contributes considerably to the transport, and the rate of oxidation exceeds that calculated using Wagner's theory. The rate of grain boundary diffusion depends on the microstructure of the oxide films formed, which is difficult to control [8]. For this reason, measured oxidation rates are often not well reproducible.

© 2007, first edition, EPFL Press

Formation of multiple layers

On certain metals, the films that form consist of two, sometimes three, different oxides:

- on iron: FeO, Fe_3O_4, Fe_2O_3
- on cobalt: CoO, Co_3O_4
- on copper: Cu_2O, CuO.

When multiple layers are formed, the growth of any one of the oxides can limit the rate of oxidation, depending conditions. The oxidation of iron offers a good illustration of this behavior.

The iron-oxygen phase diagram (Figure 9.15) shows the stability domains of the three oxides $Fe_{1-x}O$ (wustite), Fe_3O_4 (magnetite) et Fe_2O_3 (hematite). Wustite only exists at temperatures exceeding 570 °C. Below this temperature, the main product of the reaction is magnetite. The parabolic oxidation constant shows a discontinuity at this temperature (Figure 9.16).

Above 570 °C, we observe the formation of a triple oxide layer that, from the metal-oxide interface outwards, is composed of wustite, magnetite and then hematite (Figure 9.17). The oxygen ion content of the triple layer thus increases as one moves from the interior to the outside. Different oxidation reactions take place at the interfaces. At the metal-wustite interface, iron is oxidized into Fe^{2+}, whereas at the wustite-magnetite and magnetite-hematite interfaces, the Fe^{2+} is transformed into Fe^{3+}.

Wustite is a *p*-type semiconductor that grows by movement of cation vacancies and holes. The same mechanism holds for the magnetite layer, but the growth mechanism of hematite is more complicated. It probably involves simultaneous

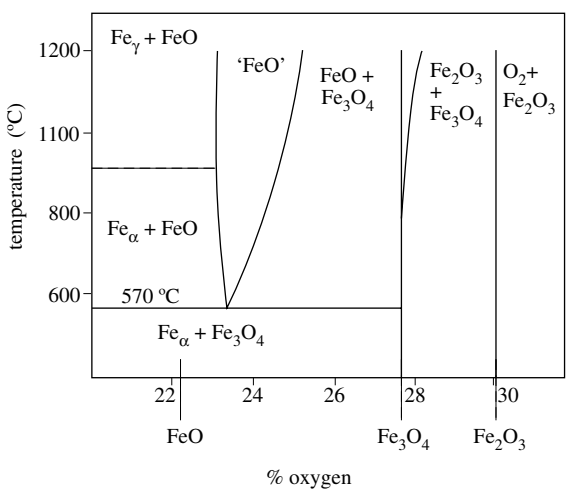


Figure 9.15 Phase diagram of Fe-O binary system [8].

© 2007, first edition, EPFL Press

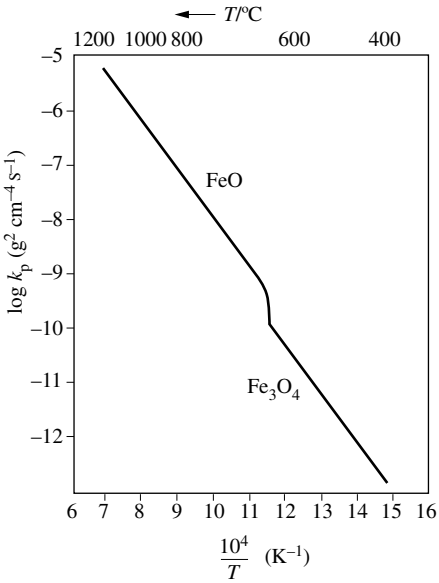


Figure 9.16 Variation of parabolic oxidation constant of iron with temperature [9].

movement of interstitial cations and oxygen vacancies. In addition, the hematite layer is porous and gaseous oxygen can find its way to the magnetite-hematite interface, where it is reduced into O^{2-}.

The thicknesses of the wustite, magnetite and hematite layers correspond to the approximate ratio of 90 to 10 to 3 [9]. Wustite therefore is the main product of the reaction. The smaller thicknesses of the magnetite and hematite layers indicate that the point defects are less mobile in these oxides. This explains why the parabolic

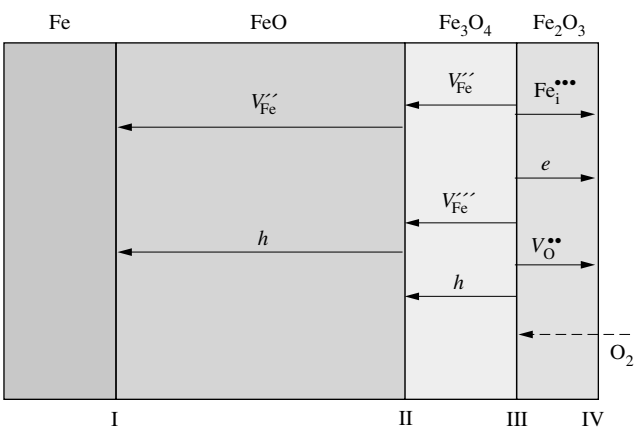


Figure 9.17 Growth mechanisms of three overlapping layers during the high-temperature oxidation of iron [9].

© 2007, first edition, EPFL Press

constant drops suddenly at 570 °C (Figure 9.16): above this temperature, the oxidation kinetics is controlled essentially by the growth of the wustite layer and, below this temperature, by that of magnetite. Contrary to the prediction of the theory of oxide growth by volume diffusion, the parabolic growth constant of wustite does not depend on the partial pressure of oxygen, because the oxygen activity at the outside surface of the wustite layer is fixed by the magnetite.

Growth stresses

Most metals used at high temperatures have a Pilling-Bedworth ratio greater than 1 ($R_{PB} > 1$). During oxidation, compressive stress therefore develops in the oxide layer. Two mechanisms act to dissipate the stress:

- plastic deformation of the oxide layer or of the underlying metal;
- cracking or spalling of the layer.

Plastic deformation

Plastic deformation of the oxide layer or of the substrate can relieve the growth-related stress and thus prevent cracking or spalling to occur. It occurs more readily if the oxide film is thin. Thick oxide layers therefore deteriorate more easily by cracking or spalling.

At high temperature, ***creep*** is the main mechanism of plastic deformation. It involves the diffusion of vacancies in the oxide or in the substrate. The vacancies preferentially coalesce at grain boundaries or at the oxide-metal interface. As a result of coalescence, microscopic cavities form that cause a weakening of the cohesion of the oxide at the grain boundaries or of the adhesion at the oxide-metal interface. This in turn facilitates cracking or spalling. Under these conditions, plastic deformation itself can thus be a source of deterioration of oxide films.

Cracking and spalling

Cracking and spalling always have a negative effect on the resistance of the metal to oxidation. In thermogravimetric experiments, the cracking of the layer is generally noticed by a sudden increase in the oxidation rate. Figure 9.18 shows the mass change of chromium samples as a function of the oxidation time for three temperatures. The discontinuities observed at 1075 °C, are obviously the result of cracking of the oxide films. At 1200 °C, the cracking starts already at the beginning of oxidation. As a consequence, no discontinuity is observed and the oxidation follows a linear rate law. In contrast, at 1000 °C, no cracks form during the entire experiment and the growth remains parabolic.

Figure 9.19 schematically shows the spalling of an oxide layer that forms during the oxidation of a metal wire. Gradually, as the reaction progresses, the diameter of the wire decreases, while the thickness of the oxide layer increases (a). As a consequence, stress relief between the layer and the substrate becomes more difficult and the layer begins to separate locally from the substrate (b). At this location, growth

© 2007, first edition, EPFL Press

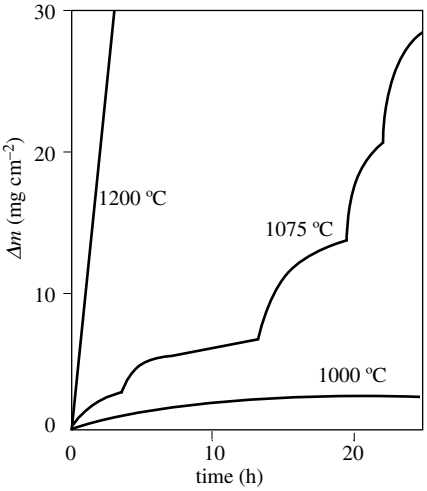


Figure 9.18 Mass increase during oxidation of Cr at different temperatures. P_{O2} = 1 bar [10].

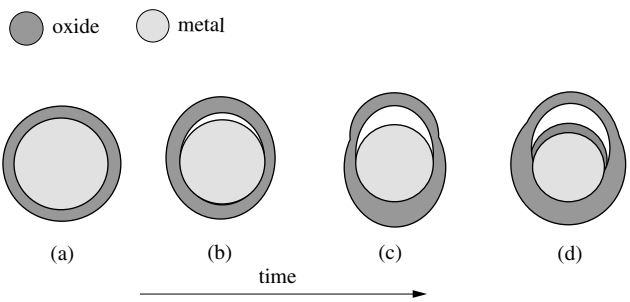


Figure 9.19 Growth of oxide layer leading to spalling (schematic): (a) adherent layer; (b) partital layer separation; (c) growth limited to adherent part of layer; and (d) formation of a new layer in the cavity.

stops because there is no more contact between the metal and the oxide (c). Sometimes, a new oxide layer, often porous, appears inside the cavity (d). The gaseous oxygen required to form this layer normally reaches it through cracks in the exterior layer. In the absence of cracks the inside wall of the outer layer can act as a source of oxygen because the oxygen activity is greater there than at the metal surface.

Growth stress generally relaxes more easily on a flat surface. The geometry of the surface and the rugosity of the substrate therefore strongly affect the adhesion of oxide layers.

9.3 OXIDATION OF ALLOYS

9.3.1 Oxidation mechanisms

The study of the oxidation behavior of binary alloys AB will allow us to highlight the most important oxidation mechanisms of multi-component alloys. The two

© 2007, first edition, EPFL Press

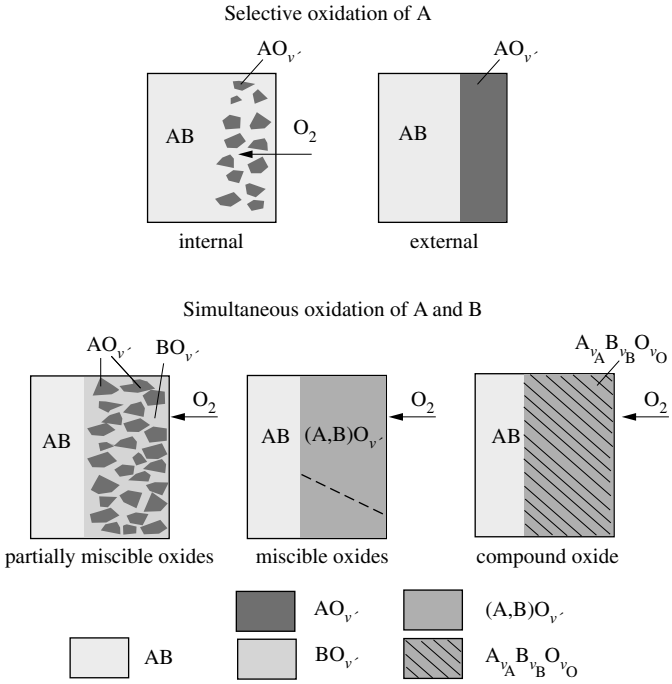


Figure 9.20 Alloy oxidation mechanisms, with the corresponding morphologies of the oxide layers.

components of a binary alloy in general do not have the same affinity for oxygen and therefore do not oxidize at the same rate. Two distinct types of behavior are observed:

- if the affinity for oxygen of A is much greater than that of B, A will undergo *selective oxidation*;
- if A and B have comparable affinities for oxygen, they undergo *simultaneous oxidation*.

Based on the observed morphology of the oxidation products, these reactions can be grouped into five mechanisms [11], sketched out in Figure 9.20**:**

- selective internal oxidation;
- selective external oxidation;
- simultaneous oxidation with formation of partially miscible oxides;
- simultaneous oxidation with formation of completely miscible oxides;
- simultaneous oxidation with formation of compound oxides.

Selective internal oxidation

Selective internal oxidation, often simply referred to as ***internal oxidation***, occurs when an alloy element A that is present in small concentration has a much greater affinity for oxygen than element B. It will then react with oxygen inside the metallic

© 2007, first edition, EPFL Press

matrix. This reaction is possible because oxygen has some solubility in metals. A precipitate $AO_{v'}$ forms whose size, typically on the order of micrometers, varies depending on reaction conditions, especially the temperature.

The rate of selective internal oxidation is limited by the rate of oxygen diffusion into the metallic matrix. With increasing oxidation time, the reaction slows because the zone in which it takes place moves towards the interior. The thickness of the two-phase layer ($B + AO_{v'}$) increases gradually as the element A is transformed into oxide. The oxidation of CoTi alloys with very low Ti content illustrates this mechanism, titanium having a much greater affinity for oxygen than cobalt.

Selective external oxidation

When an alloy contains a higher proportion of a reactive element A, the zone of internal oxidation is located closer to the surface. Furthermore, the precipitate occupies an increasing fraction of the volume. Above a critical concentration, a surface layer of oxide forms. The situation is referred to as a selective external oxidation, or more simply, ***selective oxidation***. The rate of selective external oxidation is limited by either the diffusion of ionic point defects across the oxide layer or by the diffusion of A from the interior of the alloy to the metal-oxide interface. The oxidation of Ni in NiPt alloys corresponds to this case [12].

Simultaneous oxidation with formation of partially miscible oxides

If the two components of the alloy oxidize by forming products that are only partially miscible, one observes the growth of a two-phase layer composed of the oxides $AO_{v'}$ and $BO_{v'}$. This kind of behavior can only take place within a limited range of concentrations. Figure 9.21 shows the ternary phase diagram of a system A-

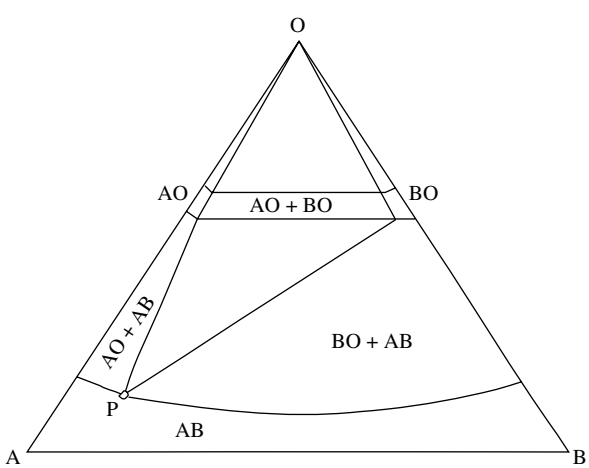


Figure 9.21 Ternary diagram A-B-O of a system forming slightly soluble oxides [11].

© 2007, first edition, EPFL Press

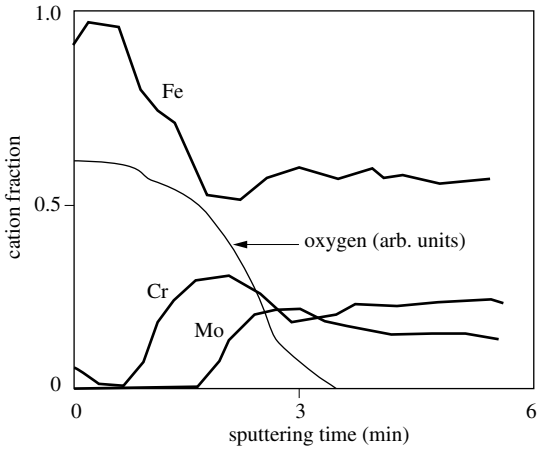


Figure 9.22 AES profiles measured on Fe-24Cr-11Mo alloy oxidized at 384 °C [14].

B-O in which $\nu' = 1$. The equilibrium between the alloy A-B and the oxides AO and BO is located at point *P*. It corresponds to the following reaction:

$$AO + B = BO + A \tag{9.55}$$

When the concentration of A becomes greater or smaller than corresponding to *P*, only the oxide AO, or respectively BO, forms. In reality, concentration gradients are established and modify the alloy composition at the oxide-metal interface. As a result, the equilibrium conditions in the reaction zone change as the reaction progresses.

The morphology of the two-phase layers formed during simultaneous oxidation does not always correspond to those presented in Figure 9.20. Often ***stratified layers*** appear. The oxidation of the alloy Cu 38% Ni (the "constantan" alloy used in thermocouples) for example, leads to the formation of an exterior layer of CuO and an interior layer of Cu_2O + NiO [13]. In the same fashion, when the alloy FeCrMo oxidizes at 400 °C, the formation of an exterior iron-oxide layer is observed, with an interior layer rich in chromium. The Auger profiles shown in Figure 9.22 illustrate this behavior: the oxidation of molybdenum is more difficult under these conditions and its concentration increases at the oxide-metal interface. The exterior layer is rich in iron oxide, which can be explained by the higher mobility of the iron ions in the oxide.

Simultaneous oxidation with formation of completely miscible oxides

If the oxides of elements A and B are miscible, the simultaneous oxidation of a binary alloy produces an oxide layer that contains a varying proportion of the cations A and B. Their stoichiometry is expressed with the formula $(A,B)_{\nu_M}O_{\nu_O}$. For example, the cubic oxides CoO, FeO, MnO, and NiO are miscible because the cations can mutually replace one another. In the same way, the trivalent oxides of aluminum and chromium form solid solutions corresponding to the formula $(Cr,Al)_2O_3$.

© 2007, first edition, EPFL Press

The stoichiometry of miscible oxides formed during alloy oxidation depends not only on the concentration of the alloy elements but also on the transport properties of the different cations in the oxide. If the cation mobility varies greatly, the stoichiometry changes with growing film thickness.

Simultaneous oxidation with formation of compound oxides

Under certain conditions, miscible oxides form stoichiometric compounds, well known examples are the oxides of the spinel type AB_2O_4. In this structure, the cations occupy well defined lattice positions and they exhibit a relatively small diffusion coefficient. Compact oxide films having the spinel structure therefore are generally beneficial for the oxidation resistance of an alloy. Unfortunately, the formation of spinels requires specific thermodynamic conditions, which often do not favor the formation of compact layers of sufficient thickness. The oxide $Fe_{1+x}Cr_{2-x}O_4$ is an exception to this rule, because its composition can vary considerably without a change in structure.

Other phenomena

The morphologies shown in Figure 9.20 are somewhat idealized. In fact, the oxide scales resulting from simultaneous oxidation of alloys often have a more complicated structure: several stacked layers, presence of precipitates in the adjacent metal zone due to internal oxidation, presence of metal islands in the oxide films, formation of pores and of cavities, etc. In addition, the oxidation of an alloy often exhibits non steady state behavior involving an initial phase during which the structure and the composition of the scales change (***transient oxidation***).

9.3.2 Kinetics of selective oxidation of binary alloys

The kinetics of selective oxidation of binary alloys depends on the relative diffusion rates of oxygen and of element A, respectively, in the metal matrix. Inspection of Figure 9.23 will allow us to understand this behavior.

The figure shows the linearized concentration profiles of dissolved oxygen (in the form of O atoms) and of metal A. At the reaction front, both concentrations are quite small, as long as the oxidation reaction is sufficiently fast. Their value corresponds to the saturation concentration of the formed oxide. The displacement rate of the reaction front lies between two limiting cases, determined by the rate controlling step of the oxidation reaction:

- oxygen diffusion (selective internal oxidation);
- diffusion of metal A (selective external oxidation).

Rate limited by oxygen diffusion

Figure 9.23 (b) shows schematically the concentration profiles when oxygen diffusion controls the rate of oxidation. The flux of dissolved oxygen, N_O, is given by

© 2007, first edition, EPFL Press

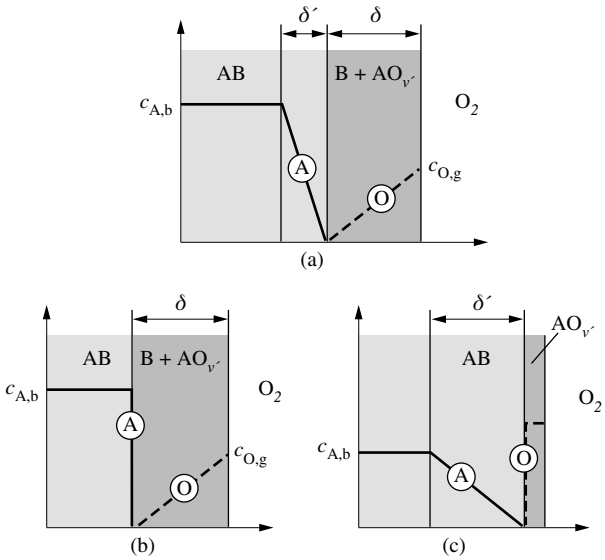


Figure 9.23 Selective oxidation of A in alloys A-B: (a) concentration profiles of A and of oxygen resulting in mixed control; (b) the oxidation rate controlled by diffusion of oxygen; and (c) oxidation rate controlled by diffusion of A.

the equation (9.56) where $c_{O,g}$ represents the dissolved oxygen concentration at the metal-gas interface and δ the thickness of the diffusion zone:

$$N_O = -D_O \frac{c_{O,g}}{\delta} \tag{9.56}$$

The number of moles of oxide $AO_{v'}$ formed per unit time and per unit of surface area is proportional to the oxygen flux.

$$\frac{dn_{AO}}{dt} = -\frac{N_O}{v'} \tag{9.57}$$

Here n_{AO} the number of moles of oxide formed per unit of surface area (mol m^{-2}). It is equal to the quantity of element A present in the layer of thickness δ.

$$n_{AO} = c_{A,b}\,\delta \tag{9.58}$$

Here, $c_{A,b}$ represents the concentration of A in the alloy, in mol m^{-3}. Differentiation of this expression with respect to time gives:

$$\frac{dn_{AO}}{dt} = c_{A,b}\frac{d\delta}{dt} \tag{9.59}$$

© 2007, first edition, EPFL Press

Equations (9.56), (9.57) and (9.59) thus lead to:

$$\frac{d\delta}{dt} = D_O \frac{c_{O,g}}{c_{A,b} v' \delta} \tag{9.60}$$

Integration with respect to time yields:

$$\delta = \left(\frac{2 D_O c_{O,g}}{v' c_{A,b}} \right)^{1/2} t^{1/2} \tag{9.61}$$

With (9.58) we find for the number of moles of oxide formed:

$$n_{AO} = \left(\frac{2 D_O \; c_{O,g} \, c_{A,b}}{v'} \right)^{1/2} t^{1/2} \tag{9.62}$$

Equation (9.62) can be used to obtain an expression for the parabolic oxidation constant, defined by (9.18). Setting $m_{AO} = M_{AO} \, n_{AO}$, where M_{AO} is the molecular mass of the oxide $AO_{v'}$ and m_{AO} is its mass, we obtain:

$$k_p = \frac{1}{v'} D_O \, c_{O,g} \, c_{A,b} \, M_{AO}^2 \tag{9.63}$$

A more rigorous mathematical treatment, based on Fick's second law, gives the same result, except for a multiplication factor.

According to equation (9.62), the amount of metal oxide after time t is proportional to the square root of the dissolved oxygen concentration at the metal-gas interface. Henry's law provides a relation between this concentration and the partial pressure of oxygen:

$$c_{O,g} = k_H P_{O_2}^{1/2} \tag{9.64}$$

The amount of oxide formed in a given time thus varies proportional to $P_{O2}^{1/4}$.

Rate limited by the diffusion of metal A

According to (9.61), the thickness of the zone corroded by internal oxidation is proportional to $(c_{O,g}/c_{A,b})^{1/2}$. If this ratio is small, in other words if the concentration of A is large and the solubility of oxygen in the alloy is small, the reaction takes place close to the surface. A compact oxide layer can then form and we have a situation of selective external corrosion.

The rate of oxidation is, in this case, controlled by the transport of element A in the metal (Figure 9.23 (c)). To calculate the flux, we assume that the concentration of A is practically zero at the metal-oxide interface. Thus, for a linear concentration profile:

$$N_A = D_A \frac{c_{A,b}}{\delta'} \tag{9.65}$$

© 2007, first edition, EPFL Press

Each mole of A that reacts at the metal-oxide interface produces a mole of the oxide $AO_{\nu'}$. The oxidation rate (mol m^{-2} s^{-1}) is therefore equal to:

$$\frac{dn_{AO}}{dt} = D_A \frac{c_{A,b}}{\delta'} \tag{9.66}$$

A simple mass balance allows us to relate the rate of oxidation dn_{AO}/dt to the thickness change of the zone depleted in element A, $d\delta'/dt$.

After a time t, n_{AO} moles of the oxide $AO_{\nu'}$ are formed, and the depleted layer attains the thickness δ'. The total number of moles of A that disappear, n_A, is equal to the number of moles of oxide formed: $n_A = n_{AO}$. If the concentration profile of metal A in the depleted layer is linear, the number of moles of A present in this layer is equal to (1/2) $c_{A,b}$ δ', where $c_{A,b}$ corresponds to the concentration of A in the bulk alloy. In the absence of reaction, a layer of the same thickness would contain $c_{A,b}$ δ' moles of A. The number of moles of A lost during the reaction is thus equal to half of the number of moles initially present.

$$n_A = c_{A,b} \frac{\delta'}{2} = n_{AO} \tag{9.67}$$

Taking the derivative with respect to time:

$$\frac{c_{A,b}}{2} \frac{d\delta'}{dt} = \frac{dn_{AO}}{dt} \tag{9.68}$$

With (9.66) we obtain the differential equation (9.69).

$$\frac{d\delta'}{dt} = \frac{2 D_A}{\delta'} \tag{9.69}$$

By integrating this expression with respect to time, we find the thickness of the layer depleted in A: $\delta' = (4 D_A t)^{1/2}$. With (9.67), and by setting $m_{AO} = n_{AO} M_{AO}$ we find the mass of oxide $AO_{\nu'}$ formed by selective oxidation:

$$m_{AO} = D_A^{1/2} c_{A,b} M_{AO} t^{1/2} \tag{9.70}$$

This equation has the same mathematical form as the parabolic growth law (9.18). The oxidation constant is equal to:

$$k_p = (1/2) D_A c_{A,b}^2 M_{AO}^2 \qquad (kg^2 m^{-4} s^{-1}) \tag{9.71}$$

The discussion shows that parabolic oxidation kinetics can be explained by different reaction mechanisms. Indeed, the equations (9.50) (9.51), (9.62) and (9.71) each give a theoretical interpretation of the parabolic oxidation constant corresponding to a different mechanism. To distinguish among these it is necessary to complement oxidation experiments by metallographic studies.

© 2007, first edition, EPFL Press

Transition criterion

The parameter γ expresses the ratio between the lengths δ' and δ that characterize respectively, the rates of selective external and internal oxidation. These lengths are given by equations (9.69) and (9.62).

$$\gamma = \frac{\delta'}{\delta} = \left(2\nu' \frac{D_A}{D_O} \frac{c_{A,b}}{c_O}\right)^{1/2} \tag{9.72}$$

The parameter γ can be used as a criterion for the transition from selective external oxidation to selective internal oxidation: if $\gamma>1$, external oxidation leads to formation of a compact layer; if $\gamma<1$, the internal oxidation occurs.

Formula (9.72) has been derived from much-simplified models, assuming linear concentration profiles and taking into account limiting cases only. However, a more detailed calculation [15] yields a comparable result. The theory presented in this section assumes a planar reaction front. Unfortunately, most often this does not correspond to reality: the reaction interface becomes rapidly unstable under conditions where the diffusion of metal A limits the rate of selective corrosion. In fact, whenever the oxide exhibits a slight protrusion into the metal, the reaction at that point accelerates because the positive curvature of the interface makes it more accessible for diffusion. Figure 9.24(a) schematically represents this situation. Grain boundary diffusion is another factor of instability (Figure 9.24 (b)), in particular at not too high temperatures. Under these conditions, oxygen diffuses preferentially through grain boundaries and these oxidize before the rest of the alloy.

In the absence of a planar reaction front, we cannot, in general, predict the rate of selective oxidation of alloys. The criterion (9.72) is nevertheless useful, because it allows us to understand qualitatively the influence of different parameters on the oxidation behavior of binary alloys.

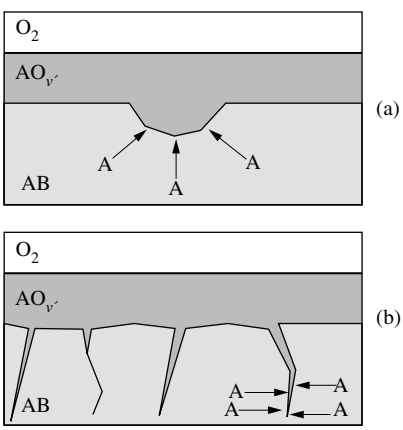


Figure 9.24 Origins of instability of oxide-metal interface: (a) preferred diffusion to protrusions; and (b) preferred oxidation at grain boundaries.

© 2007, first edition, EPFL Press

9.3.3 Kinetics of simultaneous oxidation of binary alloys

In this section we describe two distinct situations:

- the growth of miscible oxides of formula $(A,B)O_v$, containing only a small concentration of element A;
- the formation of spinel structures of formula AB_2O_4.

Miscible oxides

Most compact oxide layers grow by cation diffusion and their growth rate depends on the concentration of interstitial cations or vacancies. Let us assume that an *n*-type bivalent oxide MO contains a solute in low concentration. According to equation (9.34), if the solute is a monovalent ion A^+ (electron acceptor), it increases the concentration of interstitial cations.

$$c_{M_i^{\bullet\bullet}} = \frac{1}{2} c_{A'} \tag{9.34}$$

Because the growth rate of n-type oxides is proportional to the concentration of interstitial cations, it will increase. Inversely, if the solute is a trivalent ion D^{3+} (electron donor), the concentration of the interstitial cations decreases, and so does the rate of oxidation. Indeed, according to (9.35), donor ions create cation vacancies,

$$c_{V_M''} = \frac{1}{2} c_{D^\bullet} \tag{9.35}$$

which then react with interstitial cations thereby annihilating them:

$$M_i^{\bullet\bullet} + V_M'' = 0 \tag{9.73}$$

Table 9.25 summarizes the described effects, known as the ***Wagner-Hauffe rules***.

Table 9.25 Effect of solute valence on the parabolic rate constant k_p of a bivalent oxide.

Semiconductor type	Diffusing species	Monovalent cation solute (electron acceptor)	Trivalent cation solute (electron donor)
n	$M_i^{\bullet\bullet}$	$c_{M_i^{\bullet\bullet}}$ increases k_p increases	$c_{M_i^{\bullet\bullet}}$ decreases k_p decreases
p	V_M''	$c_{V_M''}$ increases k_p decreases	$c_{V_M''}$ increases k_p increases

The Wagner-Hauffe rules, illustrated here for a bivalent oxide, apply in fact to any oxide. Thus, in a general fashion, in those metals that form *n-type oxides*, alloying elements of lower valence will accelerate corrosion, and elements of higher valence will slow it. Inversely, metals that form p-type oxides oxidize more rapidly in presence of alloying elements of higher valence, and less rapidly when alloyed with elements of lower valence. The results presented in Table 9.26 serve as an illustration of the Wagner-Hauffe rules.

© 2007, first edition, EPFL Press

Unfortunately, the practical usefulness of these rules is limited, because they only apply to the simultaneous oxidation of binary alloys where one component is present in small concentration. Such alloys, however, most often exhibit selective oxidation.

Table 9.26 Illustration of the Wagner-Hauffe rules (adapted from [16]).

Metal/alloy	Oxide	Solute	Type	Temperature (°C)	k_p ($10^{-10}g^2cm^{-4}s^{--1}$)
Ni	NiO	–	p	1000	3.8
Ni-0.3 %Cr		Cr			15
Zn	ZnO	-	n	390	8
Zn-0.4%Li		Li^+			2000
Zn-1.0%Al		Al^{3+}			0.1

Compound oxides

At high solute concentrations, the alloy oxidation exhibits generally a complex behavior. Depending on the alloy system and the experimental conditions one observes the formation of mixed compounds, solid solutions, oxide mixtures, or numerous phases. The oxidation of ***Co-Cr alloys*** serves as an illustration.

In Figure 9.27 the parabolic oxidation constant for this alloy at 1100 °C is plotted as a function of the chromium content. Its value decreases sharply as the concentration of chromium reaches 30%. The reason is that at the temperature of the experiments chromium-rich alloys form a compact oxide film, made essentially of Cr_2O_3 , while low chromium alloys form a number of oxides (Figure 9.28). In low chromium alloys an exterior layer of CoO covers a two-phase layer of CoO and $CoCr_2O_4$. Below, a zone of internal oxidation contains a dispersion of Cr_2O_3 within the metal matrix.

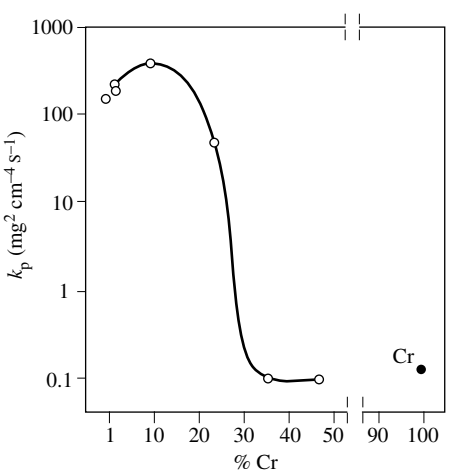


Figure 9.27 Variation of the parabolic oxidation constant of Co-Cr alloys as a function of their Cr content at 1100 °C, P_{O2} = 1 bar [13].

© 2007, first edition, EPFL Press

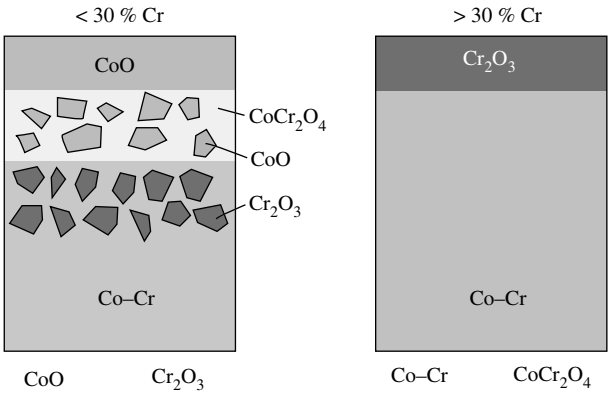


Figure 9.28 Morphology of oxide layers formed on Co-Cr alloys of different Cr content (schematic).

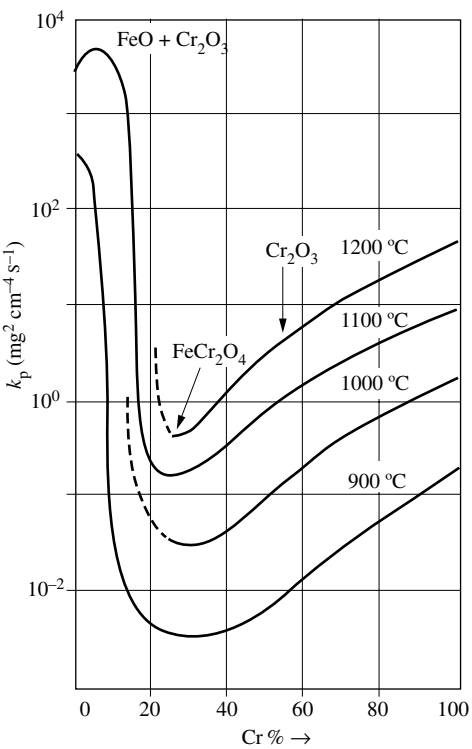


Figure 9.29 Variation of parabolic rate constant for oxidation of Fe-Cr alloys as a function of Cr concentration [17].

© 2007, first edition, EPFL Press

The formation of multiple layers can be explained in the following way: When the Cr-Co alloy is exposed to oxygen, the oxides Cr_2O_3 and CoO form. Cr_2O_3 is thermodynamically more stable, and this shifts the equilibrium of reaction (9.74) to the right.

$$2\,Cr + 3\,CoO = Cr_2O_3 + 3\,Co \tag{9.74}$$

As a result, selective oxidation of chromium is observed. The activity of chromium at the interface decreases as the reaction progresses. The equilibrium then shifts gradually towards the left, permitting thus the simultaneous existence of the two oxides. By reacting with each other they form a spinel phase:

$$CoO + Cr_2O_3 = CoCr_2O_4 \tag{9.75}$$

Oxidation of a Co-Cr alloy of low Cr content thus leads to formation of three different oxides: Cr_2O_3, CoO, and $CoCr_2O_4$.

The particular spatial distribution of the oxides (Figure 9.28) results from their growth kinetics. CoO grows more quickly than $CoCr_2O_4$, even though the mechanism is similar, i.e., based on cation diffusion. As a result of its solubility in the metal phase, oxygen can diffuse into the alloy and react with chromium, leading to the appearance of Cr_2O_3 precipitates. The discussed behavior of Co-Cr alloys brings together all the mechanisms presented in Figure 9.20.

The oxidation of Ni-Cr and Fe-Cr alloys is similar to that for Co-Cr. Figure 9.29 shows the variation of the parabolic rate constant for Fe-Cr alloys as a function of the chromium content. At the curve's minimum, which corresponds to approximately 30% chromium, a compact layer of the mixed oxide $FeCr_2O_4$ appears.

9.3.4 Multi-component alloys

Most alloys that are resistant to high-temperature corrosion are based on the ternary system M-Cr-Al, where M stands for Fe, Ni or Co. While metallic aluminum has a low melting point (660 °C), the thermodynamic stability of its oxide (melting temperature 2030 °C) exceeds even those of the oxides of Cr, Fe, Ni et Co (Figure 2.2). As a consequence, aluminum oxide grows more slowly (Figure 9.14) and formation of a compact Al_2O_3 film therefore can slow down the rate of alloy oxidation. Figure 9.30 shows, in the ternary diagram Ni-Cr-Al, the composition of layers formed at 1000°C. This kind of experimental result is often called an "oxidation map". According to Figure 9.30, the binary alloys NiAl must contain more than 15% Al for an Al_2O_3 film to form. If the concentration is lower, alloyed aluminum undergoes a selective internal oxidation.

The presence of chromium lowers this limit by about 10%. Like aluminum, chromium has strong affinity for oxygen. Its presence therefore shifts the reaction zone between oxygen and aluminum towards the surface where, because of its greater stability, Al_2O_3 forms a compact layer.

The so-called ***superalloys*** are based on the ternary system M-Cr-Al and they contain additional elements in minor quantity: Y, Ce, Zr, W, C, Ti, Mo. Some of these

© 2007, first edition, EPFL Press

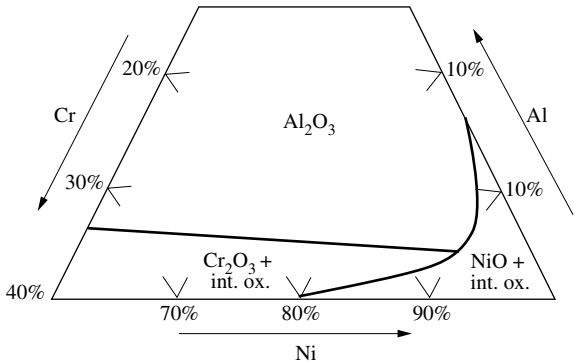


Figure 9.30 Oxidation map of Cr-Ni-Al system at 1000 °C [13].

elements (W, C, Ti) improve the mechanical properties of the alloys, in particular the creep resistance. Others (Y, Ce, Zr), called "active elements" because of their high affinity for oxygen (compared to chromium), reduce the oxidation rate. In addition, these elements improve the adhesion of the oxide layers to the substrate. This effect has been explained by an increase of the relative importance of oxygen diffusion at the expense of cation diffusion, with the result that the formation of cavities near the metal-oxide interface is suppressed. Another explanation proposes that these elements suppress sulfur segregation at the oxide-metal interface [18].

The oxidation of superalloys involves the same types of mechanisms as the oxidation of binary alloys, but the interactions between the different reactions are significantly more complicated. In addition, the structure and the composition of the oxides that form change during the course of the reaction in a manner that is difficult to predict (transient oxidation). For example, during the oxidation of the alloy U 700 (Ni-15Cr-19Co-5Mo-4.5Al-3.5Ti, trace elements) between 900 and 1000 °C, the precipitates Al_2O_3 + Cr_2O_3 form first by internal oxidation, followed by precipitates of the oxide $(Al,Cr)_2O_3$, and a layer of Al_2O_3+TiN. Finally, at steady state, one finds a layer of Cr_2O_3+$NiCr_2O_4$+TiO_2 and the precipitates Al_2O_3 +TiN +(Mo,Ni)N [13].

Coatings

Coatings can provide protection against high-temperature corrosion. Often they are applied by ***plasma spraying***. In this process, the ingredients of the coating are introduced in the form of powder into a plasma spray gun equipped with two water cooled electrodes. An electric arc between the electrodes ionizes the gas in the chamber and produces a high temperature that leads to melting of the powder. The molten powder is projected through a nozzle onto the alloy surface, where it solidifies forming a coating. Because of the high temperatures achieved in the plasma gun even refractory materials can be applied by spray coating

© 2007, first edition, EPFL Press

Another process, ***cementation***, consists of bringing the alloy into contact with a powder and then heating them together. A coating forms by interdiffusion. This method serves, notably, to apply aluminum-based protective coatings.

Coatings used to protect gas turbines against high-temperature corrosion can be:

- aluminum-based;
- chromium-based;
- silicon-based;
- CrAlY-based;
- ceramics-based.

Coatings based on MCrAlY (M = Fe, Ni, Co) are particularly interesting because they contain the same elements as the superalloys. As they are not required to sustain mechanical load, their composition can be chosen uniquely to optimize their corrosion resistance. For example, by increasing the concentration of active elements such as Al and Y one improves the adhesion of the oxide scales.

9.4 CORROSION IN THE PRESENCE OF SULFUR COMPOUNDS

9.4.1 Thermodynamics of the metal-sulfur-oxygen system

The presence of sulfur compounds (H_2S, SO_2, SO_3) in a hot gas generally promotes high temperature corrosion. To understand the phenomena in question we first study the chemical equilibria in the metal-sulfur-oxygen system.

Equilibria in the system M-S_2

Sulfur is a member of group VI of the periodic table, just below oxygen. In analogy to high temperature oxidation, metals react with gaseous sulfur to form sulfides.

$$M + \frac{\nu'}{2} S_{2(g)} = M S_{\nu'} \tag{9.76}$$

The standard free energy of this reaction varies with temperature. Figure 9.31, comparable to the Ellingham diagrams for oxide formation, provides the standard free energy for the formation of different metal sulfides, normalized according to the reaction of one mole of gaseous S_2 (ν_M and ν_S represent the stoichiometric coefficients for the metal and for sulfur, respectively):

$$2\frac{\nu_M}{\nu_S} M + S_{2(g)} = \frac{2}{\nu_S} M_{\nu_M} S_{\nu_S} \tag{9.77}$$

It follows from a comparison of Figures 9.31 and 2.2 that, from a thermodynamic point of view, the sulfides are less stable than the oxides.

© 2007, first edition, EPFL Press

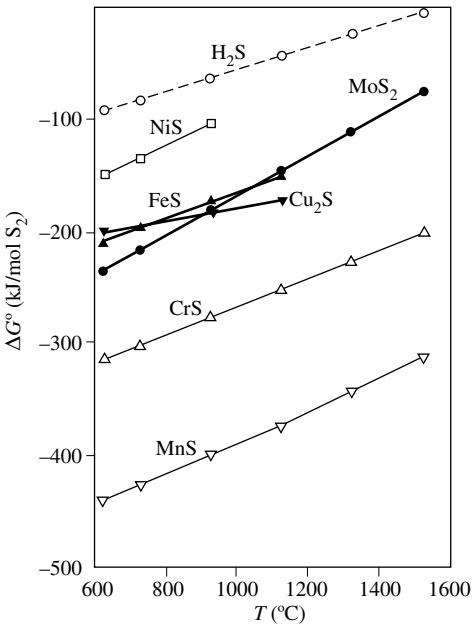


Figure 9.31 Ellingham diagram showing the variation of the standard free energy of formation of sulfides as a function of temperature. The values refer to the reaction of one mole of S_2 in the gaseous state. In order to calculate the curves, data in ref. [1] were used. The boiling temperature of sulfur is 609 °C.

Equilibria in the system M-SO_2-O_2

Under oxidizing conditions, sulfur is normally present in the form of SO_2 or SO_3. Depending on the partial pressure of oxygen, reaction with a metal leads to either oxides, sulfides or sulfates. In the following, different possible reactions of a metal leading to the bivalent compounds MO, MS or MSO_4 are listed:

$$M + \frac{1}{2} O_2 = MO \tag{9.78}$$

$$M + SO_2 = MS + O_2 \tag{9.79}$$

$$MS + 2O_2 = MSO_4 \tag{9.80}$$

$$MS + \frac{3}{2}O_2 = MO + SO_2 \tag{9.81}$$

$$MO + SO_2 + \frac{1}{2}O_2 = MSO_4 \tag{9.82}$$

© 2007, first edition, EPFL Press

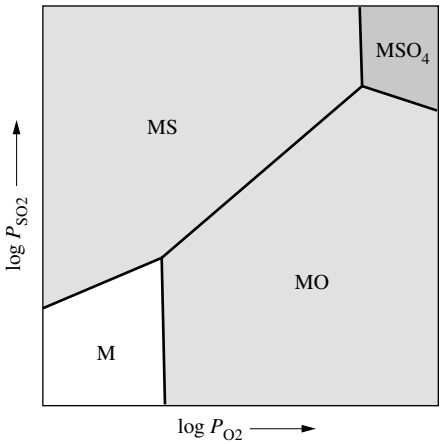


Figure 9.32 Equilibrium diagram of the system M-SO_2-O_2 showing the stability regions for different species in the presence of SO_2 and O_2 (schematic).

Figure 9.32 schematically presents the stability domains of the different solid compounds in the log P_{SO2} - log P_{O2} plane. Low partial pressures of oxygen favor the formation of sulfides, whereas a high oxygen partial pressure favors the appearance of oxides and of sulfate.

Phase stability diagrams refer to a specific temperature. In order to determine them, it is necessary to know the equilibrium constants for the reactions at that temperature. Table 9.33 presents the different equilibria represented in Figure 9.32 using the definition pK = –log K.

Table 9.33 Equilibria in the system M-O_2-SO_2.

compounds	reaction	equilibrium constant	equation
M/MO	9.78	$K_2 = P_{O_2}^{-1/2}$	$\log P_{O_2} = 2\,pK_2$
M/MS	9.79	$K_3 = P_{O_2} P_{SO_2}^{-1}$	$\log P_{SO_2} = -pK_3 + \log P_{O_2}$
MS/MSO_4	9.80	$K_4 = P_{O_2}^{-2}$	$\log P_{O_2} = 1/2\,pK_4$
MS/MO	9.81	$K_5 = P_{SO_2}\ P_{O_2}^{-3/2}$	$\log P_{SO_2} = -pK_5 + 3/2 \log P_{O_2}$
MO/MSO_4	9.82	$K_6 = P_{O_2}^{-1/2}\ P_{SO_2}^{-1}$	$\log P_{SO_2} = pK_6 - 1/2 \log P_{O_2}$

Figure 9.34 shows the phase stability diagram for nickel at 1000 K in the logP_{SO2}–logP_{O2} plane, taking into account the sulfides NiS and Ni_3S_2. The equations (9.83) and (9.84) describe the equilibria of oxygen with SO_2, SO_3 and $S_{2(g)}$.

$$SO_2 + 1/2\,O_2 = SO_3 \tag{9.83}$$

$$S_{2(g)} + 2\,O_2 = 2\,SO_2 \tag{9.84}$$

© 2007, first edition, EPFL Press

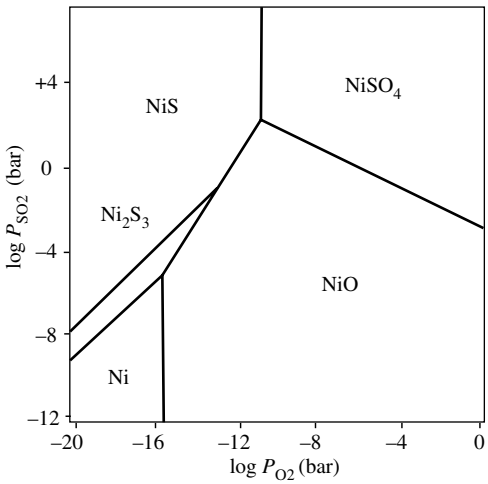


Figure 9.34 Equilibrium diagram for Ni-SO_2-O_2 at 727 °C [20].

According to these reaction equilibria every point in the logP_{SO2}-logP_{O2} plane corresponds to a certain partial pressure of SO_3 and S_2. We can therefore also draw the phase stability diagram in the planes logP_{SO3}-logP_{O2} or logP_{S2}-logP_{O2}.

Phase stability diagrams present, in graphical form, the thermodynamic data for the corrosion of metals in presence of sulfur and oxygen at a given temperature. In this sense, they are comparable to the *E*–pH diagrams that present the thermodynamic data for aqueous corrosion.

9.4.2 Sulfidation

The high-temperature reaction (9.76) between a metal and gaseous sulfur is called ***sulfidation***. It produces solid or liquid sulfides. The free energy of formation of sulfides is inferior to that of oxides of same stoichiometry (Table 9.35).

Table 9.35 Standard free energies of formation of sulfides and oxides at 1000 K [1].

Sulfide	$\Delta G°$ / kJ mole^{-1}	Oxide	$\Delta G°$ / kJ mole^{-1}
NiS	–67	NiO	–149
FeS	–98	FeO	–207
MnS	–213	MnO	–312
Al_2S_3	–508	Al_2O_3	–1356
ZrS_2	–522	ZrO_2	–908

© 2007, first edition, EPFL Press

A lower energy of formation means that sulfides have smaller binding energies than corresponding oxides. This has two consequences that affect corrosion:

- for the same temperature, the number of point defects in sulfides is higher than in oxides;
- the melting temperature of sulfides is lower than that of oxides.

Sulfides contain a particularly high *number of point defects* because of their variable stoichiometry, which depends on the forming conditions. In the solid state, the concentration of point defects can reach values on the order of 1 %. This greatly facilitates diffusion and allows compact sulfide scales to grow faster than oxide films. Table 9.36 compares the parabolic growth constants of sulfides and oxides for Ni, Co, Fe and Cr. For the same temperature, the values for the sulfides are 2 to 7 orders of magnitude greater than those of the oxides.

Table 9.36 Parabolic rate constants of oxidation and of sulfidation [21].

		Ni	Co	Fe	Cr
O_2	k_p ($g^2\ cm^{-4}\ s^{-1}$)	9.1×10^{-11}	1.6×10^{-9}	5.5×10^{-8}	4.5×10^{-12}
	T (°C)	1000	950	800	1000
S_2	k_p ($g^2\ cm^{-4}\ s^{-1}$)	8.5×10^{-4}	6.7×10^{-6}	8.1×10^{-6}	8.1×10^{-7}
	T (°C)	650	720	800	1000

The *melting temperature* of sulfides is generally lower than that of oxides. For example, for NiS it is 797 °C , compared to 1990 °C for NiO. In addition, the sulfides of Ni, Co, Fe and Cr can form eutectics with the base metal lowering even more their melting temperature: for example, in the system Ni-S the eutectic melts at 645 °C already. Nickel sulfides therefore melt easily under typical high-temperature corrosion conditions. On the other hand, chromium sulfides melt less easily; the Cr-S eutectic having a melting temperature of 1350 °C.

9.4.3 Hot corrosion

Hot corrosion refers to high-temperature corrosion that takes place in presence of molten salt on the metal surface. Sodium sulfate deposits are most common. They are the result of the reaction equilibria (9.78) -– (9.82) or of a reaction between sodium chloride and SO_3:

$$2\,NaCl + SO_3 + H_2O = Na_2SO_4 + 2\,HCl_{(g)} \tag{9.85}$$

The melting point of pure Na_2SO_4 is 884 °C, but in the presence of other salts or of certain oxides it can be lower. For example, when SO_2 is present, hot corrosion of nickel is already observed at temperatures near 700 °C, because $NiSO_4$ forms an eutectic with Na_2SO_4 that melts at 671 °C. Such low melting temperatures explain why hot corrosion is a serious threat under typical high-temperature operating conditions.

Figure 9.37 illustrates the effect of molten sodium sulfate on the rate of nickel corrosion. The results show the mass change of a nickel sample, exposed to oxygen

© 2007, first edition, EPFL Press

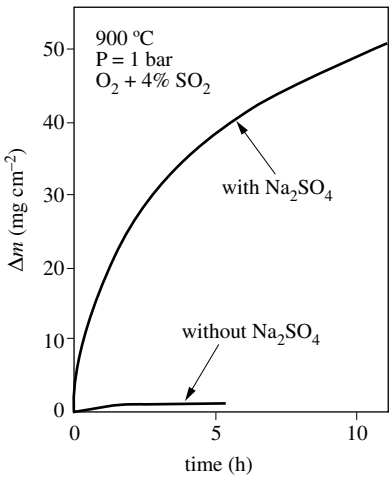


Figure 9.37 Effect of Na_2SO_4 on the rate of nickel oxidation at 900 °C, O_2 + 4% SO_2, P = 1 bar [21].

and SO_2 at 900 °C, with and without a layer of sodium sulfate on the surface. The rate of corrosion in the presence of Na_2SO_4 is much higher, indicating that the compact oxide layer, whose growth normally controls the rate of nickel oxidation, does not develop in the usual way. In fact, the molten sodium sulfate dissolves the oxide layer, a phenomenon referred to as ***fluxing***.

Similar behavior is observed with superalloys, but the mechanisms are more complicated because the surface composition changes with time due to the selective oxidation of alloy components. Figure 9.38 shows the mass variation as a function of

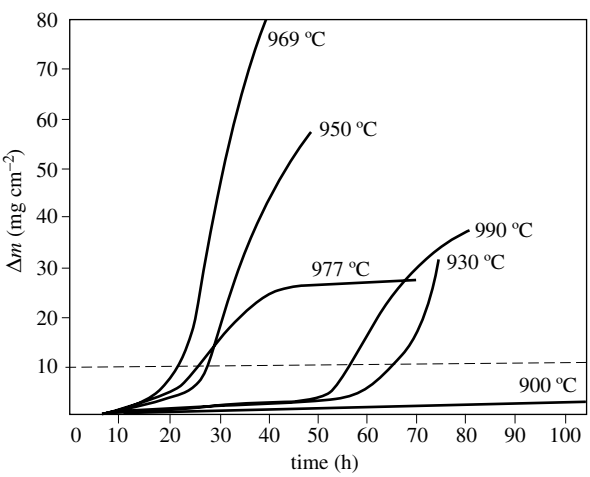


Figure 9.38 Effect of temperature on initiation time for hot corrosion of a Ni-Cr alloy (IN-738). P_{O2} = 1 bar, 1 mg cm^{-2} Na_2SO_4 [22].

© 2007, first edition, EPFL Press

oxidation time for a nickel alloy (IN738) exposed to oxygen at different temperatures. Previously, a small quantity of Na_2SO_4 (1 mg cm^{-2}) was deposited on the alloy surface to form a flux. The results show two distinct reaction periods: during the initial period the mass change remains small, but after some hours it rapidly increases. The duration of the initial period increases with decreasing temperature and at 900 °C it covers the whole length of the experiment. The ***induction period*** corresponds to the time required to form the molten eutectic and to dissolve the oxide. This process is usually faster at higher temperature.

Solubility of oxides in molten sulfate

The presence of a molten salts favors the dissolution of the surface oxides that otherwise limit the rate of high temperature corrosion. The dissolution of nickel oxide in molten sodium sulfate provides a good example. The solubility of NiO in Na_2SO_4, at a given temperature, depends on the activity of Na_2O, which is established by the equilibrium:

$$Na_2SO_4 = Na_2O + SO_3 \tag{9.86}$$

or

$$SO_4^{2-} = O^{2-} + SO_3 \tag{9.87}$$

The activity of Na_2O thus characterizes that of O^{2-} ions. A molten salt whose Na_2O activity is high is called a ***basic flux***; and one with low Na_2O activity is called an ***acidic flux***. This terminology is based on the fact that Na_2O with water forms sodium hydroxide whereas SO_3 forms sulfuric acid.

Figure 9.39 shows the solubility of NiO in molten sodium sulfate as a function of the Na_2O activity. The solubility passes through a minimum, because the stoichiometry of dissolution is not the same in acidic and basic melts.

In a basic flux, nickel oxide in the presence of oxygen is in equilibrium with the trivalent nickel species NiO_2^- :

$$NiO_{(s)} + \frac{1}{2}O^{2-} + \frac{1}{4}O_2 = NiO_2^- \tag{9.88}$$

This reaction equilibrium determines the saturation concentration of NiO_2^- ions in presence of solid NiO. The activity (concentration) of NiO_2^- ions in the molten salt depends on the partial pressure of oxygen and on the activity of the O^{2-} ions:

$$a_{NiO_2^-} \propto P_{O_2}^{1/4}\, a_{O^{2-}}^{1/2} \tag{9.89}$$

from which we get

$$\frac{d \log a_{NiO_2^-}}{d \log a_{O^{2-}}} = \frac{d \log a_{NiO_2^-}}{d \log a_{Na_2O}} = \frac{1}{2} \tag{9.90}$$

© 2007, first edition, EPFL Press

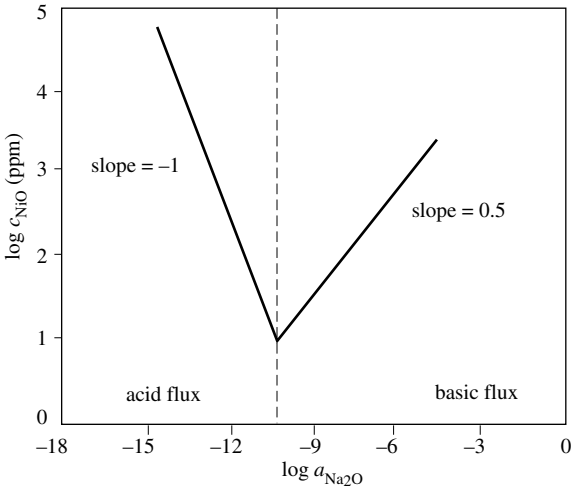


Figure 9.39 Variation of the concentration (ppm) of dissolved nickel oxide, as NiO_2^- or Ni^{2+} in molten sodium sulfate as a function of the Na_2O activity (standard state: pure Na_2O) and in the presence of oxygen. $T = 1200$ K [24].

In an acidic flux, NiO dissolution produces Ni^{2+}ions according to the equilibrium:

$$NiO_{(s)} = Ni^{2+} + O^{2-} \tag{9.91}$$

From this it follows that the activity (concentration) of Ni^{2+} ions at saturation is inversely proportional to the activity of the O^{2-} ions:

$$a_{Ni^{2+}} \propto \frac{1}{a_{O^{2-}}} \tag{9.92}$$

We thus get:

$$\frac{d \log a_{Ni^{2+}}}{d \log a_{Na_2O}} = -1 \tag{9.93}$$

The slopes of the logarithmic lines shown in Figure 9.39 are in agreement with equations (9.90) and (9.93).

Mechanism of hot corrosion of nickel

When a negative solubility gradient exists from the metal surface to the outside of the salt layer, dissolution of the NiO film covering the metal is followed by re-precipitation of NiO at some distance away from the surface. The oxide precipitate is highly porous and therefore non-protective [23]. Figure 9.40 schematically shows such a situation.

© 2007, first edition, EPFL Press

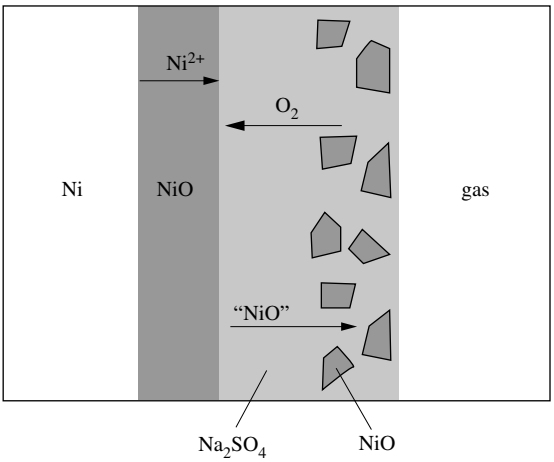


Figure 9.40 Mechanism of hot corrosion of Ni involving fluxing of oxides in molten sodium sulfate.

In order for the solubility to be highest at the oxide-salt interface, basic conditions must prevail there. The reduction of oxygen on the metal can provide an explanation:

$$O_2 + 4e = 2O^{2-} \tag{9.94}$$

However, other cathodic partial reduction reactions might also contribute to create basic conditions.

Another mechanism can also explain the hot corrosion of nickel. It takes into account the observation that the molten layers formed during hot corrosion usually

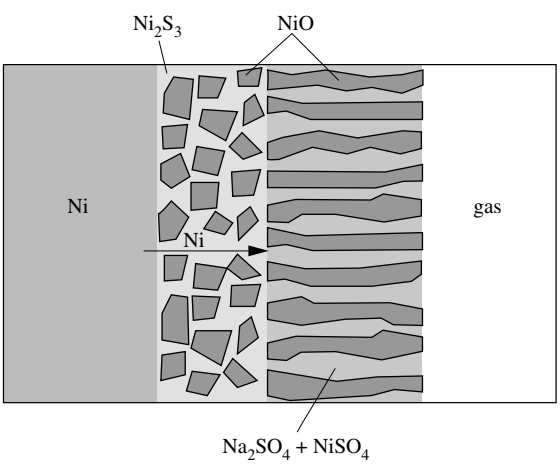


Figure 9.41 Mechanism of hot corrosion of nickel involving formation of molten nickel sulfides.

© 2007, first edition, EPFL Press

contain sulfur near the metal surface. This presence has been explained by a chemical reaction between metallic nickel and sulfates, which are thereby reduced to sulfides [21]. The mechanism is schematically shown in Figure 9.41. The sulfate ions reach the metal surface by diffusion at the oxide grain boundaries or through cracks. Upon contact with the metal surface, they are reduced to sulfide by oxidizing the nickel. This leads to the formation of liquid nickel sulfide, which partially replaces the oxide at the metal surface. Because metallic nickel has a finite solubility in molten sulfides, it diffuses towards the sulfide-sulfate interface where it reacts with oxygen, forming a non-compact oxide.

Hot corrosion due to volatile oxides

Certain volatile oxides accelerate corrosion. Vanadium pentoxide V_2O_5, for example, is produced during the combustion of petroleum-based fuels that contain organic vanadium compounds. The melting point of V_2O_5 is 690 °C. In the presence of sodium sulfate, sodium vanadates are formed:

$$V_2O_5 + Na_2SO_4 = 2\,NaVO_3 + SO_3 \tag{9.95}$$

Equation (9.95) represents the reaction in a simplified manner, because, in general, the sodium vanadates are non-stoichiometric, corresponding to the formula $(Na_2O)_{v'}V_2O_5$ where v' can take on a range of values. Nevertheless, the stoichiometric equation (9.95) allows us to highlight the role of SO_3. Specifically, the reaction takes place more easily as the partial pressure of SO_3 in the gas mixture decreases.

The system Na_2O-V_2O_5 includes a eutectic (between 35 and 40 % molar Na_2O) that melts at 535 °C. In the presence of Na_2SO_4 the melting point is even lower. The liquids that form cause fluxing of the oxide layer. This in turn leads to strong acceleration of corrosion, sometimes called ***catastrophic corrosion***.

For similar reasons, other oxides with low melting points, in particular PbO and MoO_3, reduce the corrosion resistance of alloys in high-temperature service. Generally, these oxides pose fewer practical problems than vanadium, however.

Protection methods

In order to avoid hot corrosion, sulfur and vanadium have to be eliminated from the fuel, which is, in general, not possible. Nevertheless, the effect of vanadium can be diminished by reducing the supply of oxygen into the combustion chamber. Under these conditions vanadium oxides of lower valence form, which are less dangerous. Another tactic is to add magnesium salts in order to favor the formation of magnesium vanadates. These compounds have melting temperatures that greatly exceed those of sodium vanadates.

The most common methods employed to combat hot corrosion are the optimization of the alloy composition and the application of coatings. Typically, this means to increase the chromium and aluminum content of superalloys, and to add active elements (Section 9.3).

© 2007, first edition, EPFL Press

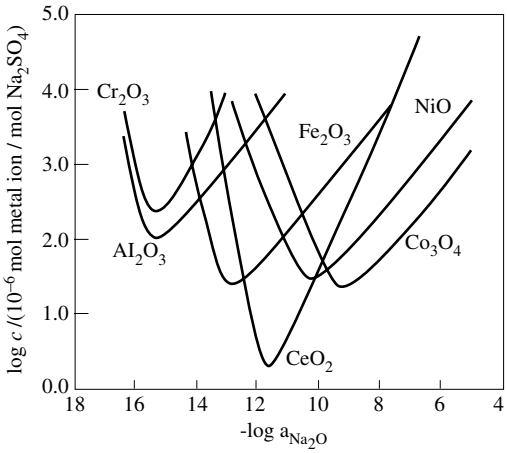


Figure 9.42 Solubility of different oxides in molten sodium sulfate as a function of melt basicity, expressed as $p_{Na2O} = -\log a_{Na2O}$; $T = 1200$ K. After [25] .

In developing suitable alloy compositions the solubility of metal oxides in sodium sulfate melts can serve as a guideline. Indeed, if hot corrosion is the result of a negative solubility gradient of metal oxides, one should try to design alloys that avoid this phenomenon under suitable operating conditions. Figure 9.42 indicates the solubility of the oxides of different metals in sodium sulfate melts at 1200 K [25]. In melts with a Na_2O activity in a range of $a_{Na2O} \approx 10^{-10}$ to 10^{-12} , the oxides of aluminum and chromium exhibit acid dissolution and therefore a positive solubility gradient.

Bibliography

P. Kofstad, *High Temperature Corrosion*, Elsevier Applied Science, London (1988) 558 pp.

N. Birks, G. H. Meier, *Introduction to High Temperature Oxidation of Metals*, Edward Arnold Publishers, London (1983) 198 pp.

M. Spiegel, "Molten Salt Induced Corrosion of Metals," in *Encyclopedia of Electrochemistry,* vol. 4, M. Stratmann, G. Fraankel, eds., Wiley-VCH (2003) pp. 597-622.

References

[1] I. Barin, *Thermochemical Data of Pure Substances*, vols. I and II, Verlag Chemie Weinheim (1989) 1737 pp.

[2] F. Armand, J. H. Davidson, *Les aciers inoxydables*, P. Lacombe, B. Baroux, G. Béranger eds., Les Editions de Physique, Les Ulis, France (1990) p. 483.

[3] S. A. Bradford, *Metals Handbook*, 9th ed., vol. 13, ASM Metals, Park Ohio (1987) p. 61.

© 2007, first edition, EPFL Press

[4] C. Wagner, *Z. phys. Chem*. B21, 25 (1933).
[5] A. T. Fromhold, *Theory of Oxidation of Metals*, vol. I, North Holland Publishing, Amsterdam (1976) pp. 108-120.
[6] F. A. Elfraie, A. Manolescu, W. W. Smeltzer, *J. Electrochem. Soc.* 132, 2489 (1985).
[7] N. Birks , G. H. Meier, *Introduction to High Temperature Oxidation of Metals*, Edward Arnold Publishers, London (1983) pp. 54, 73.
[8] H. V. Atkinson, *Oxidation of Metals,* vol. 24, 177 (1985).
[9] A. Rahmel, W. Schwenk, *Korrosion und Korrosionschutz von Stählen*, Verlag Chemie, Weinheim (1977) pp. 194, 197.
[10] P. Kofstad, *High Temperature Corrosion*, Elsevier Applied Science, London (1988) p. 235.
[11] D. P. Whittle, *High Temperature Corrosion*, NACE-6, R. A. Rapp, ed., Natl. Assoc. Corrosion Engineers, Houston Texas (1983) p. 171.
[12] C. Wagner, *J. Electrochem. Soc.*, 99, 369 (1952).
[13] P. Kofstad, High Temperature Corrosion, Elsevier Applied Science London 1988, pp. 360, 372.
[14] H. J. Mathieu, D. Landolt, *Corr. Sci*. 26, 547 (1986).
[15] C. Wagner, *Z. Elektrochemie* 63, 772 (1959).
[16] M. Fontana, N. D. Green, *Corrosion Engineering*, MacGraw Hill, New York (1967) pp. 358–359.
[17] J. Bénard, *Oxydation des métaux*, tome II, Gauthier-Villars, Paris (1962) p. 150.
[18] P. Y. Hou, J. Stringer, *Oxidation of Metals* 38, 323 (1992).
[19] P. Kofstad, *High Temperature Corrosion*, Elsevier Applied Science, London (1988) p. 372.
[20] R. D. Pehlke, *Unit Processes of Extractive Metallurgy*, Elsevier, New York (1974) p. 9.
[21] P. Kofstad, *High Temperature Corrosion*, Elsevier Applied Science, London (1988) pp. 429, 475.
[22] N. Birks, G. H. Meier, *Introduction to High Temperature Oxidation of Metals*, Edward Arnold Publishers, London (1983) p. 147.
[23] D. Shores, *High Temperature Corrosion*, R. A. Rapp, ed., NACE-6, Natl. Assoc. Corrosion Engineers, Houston Texas (1983) p. 493.
[24] D. K. Gupta, R. A. Rapp, *J. Electrochem. Soc*. 127, 2194 (1980).
[25] R. A. Rapp, *Corr. Sci.* 44, 209 (2002).

© 2007, first edition, EPFL Press

CHAPTER 10

TRIBOCORROSION

10.1 INTRODUCTION

The field of tribology is concerned with the study of friction, lubrication and wear of interacting surfaces that are in relative motion. Its name comes from the Greek "tribo" meaning to rub. ***Wear*** is a material loss resulting from mechanical interactions in a tribological contact. In presence of a corrosive environment, synergistic or antagonistic effects between wear and corrosion accelerate or slow down the rate of material loss. The terms *corrosion accelerated wear* and *wear accelerated corrosion* are sometimes used to express the accelerating effect of corrosion on wear or vice versa. In the following, the combined action of corrosion and wear on a material will be referred to as ***tribocorrosion***.

Tribocorrosion phenomena play a role in many engineering fields. They often limit the performance and the life-time of machines and devices. In addition, they can affect the safety of process installations and engineering structures and they can have implications for human well-being and health. Finally, they affect the effectiveness of certain metal shaping processes. Examples include tribocorrosion of safety systems in nuclear reactors; erosion corrosion of pipes, valves and pumps in the petroleum industry; corrosive wear of waste incinerators; wear of medical implants resulting in tissue inflammation and rejection; chemical-mechanical polishing of wavers in integrated circuit manufacturing; the action of lubricants in metal grinding and cutting.

Table 10.1 schematically presents different types of tribological contacts. They differ in terms of the contacting phases, the nature of the mechanical interactions and the presence or not of corrosion. Two broad categories of mechanical interactions are distinguished:

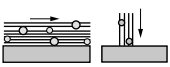
- sliding and rolling between solids;
- impingement of particles or fluids.

Wear due to sliding and rolling between solids

Friction between rubbing surfaces dissipates energy in the form of heat and causes material loss by wear. In Table 10.1 we distinguish three types of tribological contacts between solid surfaces: sliding, rolling and fretting.

© 2007, first edition, EPFL Press

Table 10.1 Different wear and tribocorrosion phenomena.

	Contact	Environment	Name
sliding or rolling between solids		inert corrosive	sliding wear corrosive wear
		inert corrosive	rolling wear corrosive wear
		inert corrosive	fretting wear fretting corrosion
impacting solids or fluids		inert corrosive	erosive wear (erosion) erosion corrosion
		inert corrosive	cavitational wear cavitation corrosion

Whenever the surfaces of two solids slide over one another, whether or not there are abrasive particles or lubricants between them, ***sliding wear*** may occur. This type of wear is frequent in machines and equipment with moving elements. In a combustion engine, for example, the movement of the piston along the cylinder wall may cause sliding wear. When the surfaces of two contacting parts slide directly on each other, we speak of a *two-body contact*. Many wear tests, as we shall see later, involve two body contacts. In practice, the sliding surfaces are often separated by a lubricant, by wear particles or by abrasive powders added on purpose. These form a so-called ***third body*** that modifies greatly the mechanical conditions prevailing in the contact, and therefore its friction and wear behavior. If particles are present between two sliding surfaces, we speak of a *three-body contact*.

Rolling wear is typically observed in bearings involving spheres or cylinders contained in a casing. Most rolling contacts are lubricated and therefore exhibit low friction and wear.

Small amplitude oscillatory motions such as vibrations in a machine can lead to ***fretting wear***. At sufficiently small amplitudes the shear forces acting on the contacting surfaces are accommodated by elastic deformation. In that case there is little or no energy dissipation and furthermore, corrosion is unimportant because no fresh surface is exposed to the environment. However, most often fretting contacts entail some sliding. In that case, there is no well-defined limit between sliding wear and fretting wear. Fretting typically involves displacement amplitudes of up to some hundred micrometers, whereas sliding contacts exhibit either oscillatory displacements of larger amplitude or a unidirectional motion. The differences in displacement mode can affect the wear behavior: at the small amplitudes typical for fretting, the wear debris are less readily evacuated from the contact zone than in sliding contacts.

© 2007, first edition, EPFL Press

Depending on their mechanical properties, they then act as an abrasive or as a lubricant. Furthermore, fretting causes repeated elastic and plastic deformation of the asperities present on the contacting surfaces, thus favoring material degradation by fatigue. ***Fretting corrosion*** designates the material loss that is due to the combined action of corrosion and wear in a fretting contact.

Wear due to impingement of particles and of fluids

Wear due to impingement of particles or fluids is called ***erosive wear*** or simply ***erosion***. Table 10.1 lists three typical situations:

- erosive wear due to the action of turbulent two phase flow parallel to a wall;
- erosive wear due to particles or fluids impacting on a wall;
- cavitational wear due to implosion of cavitation bubbles.

Erosive wear occurs when solid particles, carried by a liquid or a gas, strike a surface. Erosive wear is not always damaging, an interesting example being sand blasting, which is a technique for cleaning metal surfaces from oxide scales by projecting sand particles on them to bring about controlled wear. The impact of water droplets or of a water jet against a wall can also cause erosive wear, sometimes referred to as *impingement wear*.

The simultaneous action of erosive wear and corrosion is called ***erosion corrosion***. Erosion corrosion is often encountered in pumps and pipes exposed to turbulent flow in the presence of suspended particles. It also occurs in other situations, for example in incinerators due to ash particles that are entrained by the exhaust gas. The damage caused by the combined effect of the impingement of a liquid and of corrosion is sometimes referred to as *impingement corrosion* or *impingement attack*.

Low-pressure zones in a fluid system under turbulent flow conditions are at the origin of ***cavitational wear***. When the hydrodynamic pressure at some location temporarily drops below the vapor pressure of the fluid a fraction of the liquid is transformed into gas bubbles. The recurrent formation and implosion of these bubbles leads to material damage by local fatigue. If simultaneously corrosion takes place we speak of ***cavitation corrosion***.

The terminology used in the literature to describe tribocorrosion phenomena is not uniform and sometimes confusing. In part this is due to historical reasons, in part to the fact that many underlying mechanisms are still insufficiently understood. In the following we shall stick to the terms defined in this section.

10.2 CHEMICAL EFFECTS IN FRICTION AND LUBRICATION

10.2.1 Friction in non-lubricated contacts

Friction coefficient

To keep an object sliding at a constant speed over a flat surface, one needs to apply a force that is equal to the friction force F, which acts parallel to the surface, in

© 2007, first edition, EPFL Press

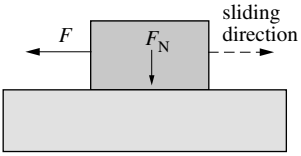


Figure 10.2 Sliding of a solid object on a plane surface.

the opposite direction (Figure 10.2). Its magnitude is proportional to the normal force F_N , perpendicular to the contact plane.

$$F = f F_N \tag{10.1}$$

The proportionality factor f is called the ***friction coefficient***. Its value depends on mechanical parameters and the presence or not of a lubricant; on chemical parameters such as humidity, oxide films, chemical additives; and on materials parameters such as hardness, ductility and microstructure.

Table 10.3 indicates the order of magnitude of the friction coefficient for different material combinations and environments. Most often in tribological systems, one seeks to reduce friction in order to reduce energy dissipation and wear. This can be achieved through the use of suitable material combinations or with liquid or solid lubricants. A number of tribological systems, however, require a high friction coefficient in combination with low wear or high wear with low friction. For example, for car brakes one wants high friction and low wear. In mechanical machining, on the other hand, one aims at a high material removal rate and low frictional heating. The examples illustrate that the goal is not always to minimize friction and wear, but rather to optimize each of these processes for a given purpose.

Table 10.3 Typical values of friction coefficients.

Contact	Environment	f
metal-metal	vacuum	> 1
metal-metal	air	0.2 – 1
metal-metal	oil lubricant	0.02 – 0.1
metal-PTFE	air	0.04 – 0.1

Friction transforms mechanical energy into heat by different mechanisms. To understand these we must remember that real surfaces are not flat but always contain some rugosity (Chapter 3). When two rough surfaces rub on each other a number of phenomena may contribute to friction (Figure 10.4):

- formation and breaking of adhesive junctions at the points of contact;
- elasto-plastic deformation of the asperities;
- material displacement by plowing.

© 2007, first edition, EPFL Press

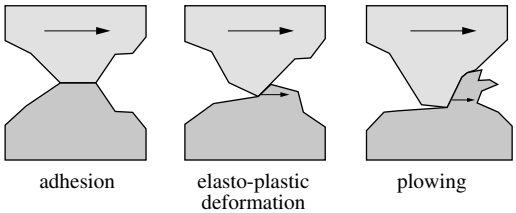


Figure 10.4 Friction mechanisms.

Formation and breaking of adhesive junctions

Figure 10.5 represents a contact between two ductile metals. Because of the surface roughness, only the asperities are touching. The ***contact area*** A_r is equal to the sum of all the asperity contact areas $A_{r,j}$:

$$A_r = \sum A_{r,j} \tag{10.2}$$

The contact area A_r is always smaller than the geometrical surface area A. The pressure acting on the contacting asperities is equal to

$$P_{asp} = \frac{F_N}{A_r} \tag{10.3}$$

If P_{asp} exceeds the elastic limit of the material, the asperities undergo plastic deformation. As a consequence, the real contact area increases until the value of F_N/A_r becomes equal to the elastic limit of the material. This is the same situation as in a hardness test: the application of a given normal force produces an indentation whose area characterizes the hardness H of the material. We can therefore write:

$$H = \frac{F_N}{A_r} \tag{10.4}$$

Sliding continuously creates and breaks asperity junctions having a shear strength τ equal to:

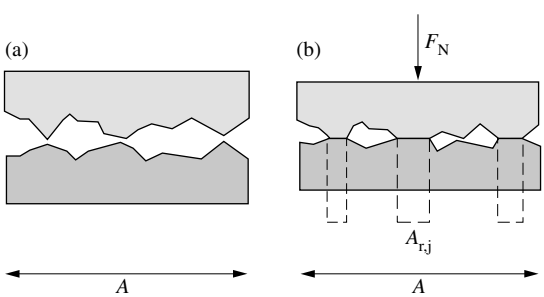


Figure 10.5 Contact of two rough surfaces of ductile materials.

© 2007, first edition, EPFL Press

$$\tau = \frac{F}{A_r} \tag{10.5}$$

Equations (10.1), (10.4) and (10.5) allow us to obtain a new expression of the friction coefficient:

$$f = \frac{F}{F_N} = \frac{\tau}{H} \tag{10.6}$$

Here, τ represents the shear strength of the junction, and H characterizes the hardness of the less hard material, implying that the harder material does not undergo deformation. The value of τ is usually taken as that of the less hard material, assuming that the breaking of junctions takes place in the softer material rather than at the contact interface.

The classical theory of friction explains a certain number of observations, in particular that the friction coefficient varies little from one metal to another. In fact, according to (10.6), the friction coefficient only depends on the ratio τ/H (and not on the absolute value of these quantities), which turns out to vary little among different metals. Equation (10.6) also allows us to understand a behavior that was observed

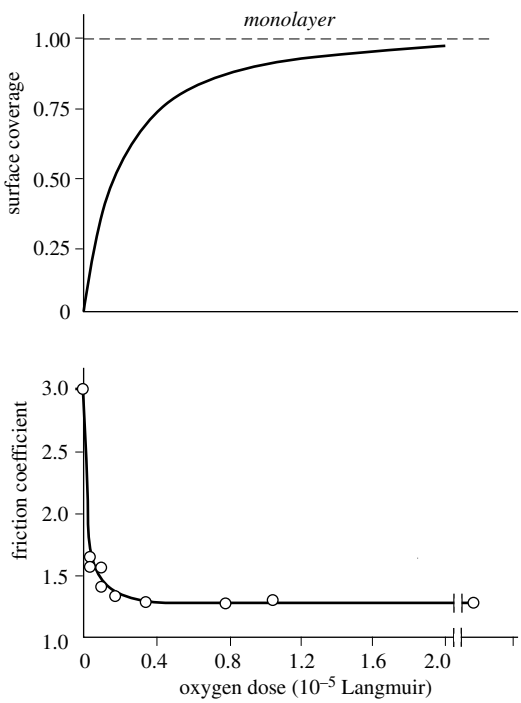


Figure 10.6 Effect of adsorbed oxygen on the friction coefficient between two tungsten surfaces: variation of the friciton coefficient as a function of oxygen dose. Pressure = 3×10^{-8} Pa. Also shown is the variation of the surface coverage as a function of the oxygen dose [1].

© 2007, first edition, EPFL Press

during the very first studies of friction, namely that the friction coefficient does not depend on the geometrical area A of contact.

However, a number of other observations are not in agreement with the described theory. First of all, the expression (10.6) yields a value of f that is too small. In fact, the hardness of metals H is approximately equal to three times the yield strength measured in uniaxial traction ($H \approx 3\sigma_e$), and their shear strength corresponds to roughly half of that value ($\tau \approx 0.5\ \sigma_e$). With (10.6), we therefore obtain a friction coefficient on the order of $f = 0.17$, which is significantly smaller than the experimental values measured for non-lubricated contacts between non-oxidized metal surfaces (Table 10.3).

The major flaw of the presented theory of friction is that it does not take into account the role of the environment. In fact, pure metal surfaces cannot exist in contact with the atmosphere, they immediately adsorb gas molecules and oxidize (Chap. 3). The presence of adsorbed species and thin oxide films prevents the formation of metal-metal bonds between the contacting surfaces. In air, the shear stress of the asperity junctions corresponds to that of oxidized surfaces rather than metal-metal contacts. The value of the shear strength in equation (10.6), and hence also that of the friction coefficient, depends therefore strongly on the chemical state of the contacting surfaces.

Figure 10.6 illustrates how the presence of adsorbed oxygen changes the friction coefficient between two tungsten surfaces [1]. In the absence of oxygen, in ultrahigh vacuum, the friction coefficient is close to $f = 3$. When oxygen is admitted, it decreases to about 1.3.

Elasto-plastic deformation

According to the classical theory of friction, the surface roughness should not influence the friction coefficient, because the asperities will deform in such a way that the true contact area depends only on the ratio of normal force over hardness. In reality, at the start of a friction experiment, one often observes a ***run-in period*** during which the friction coefficient and the surface roughness decrease. The behavior is due to an elasto-plastic deformation of the asperities in the direction of motion as shown in Figure 10.4b. The corresponding energy of deformation contributes to the energy dissipated by friction during run-in, and therefore the friction coefficient is higher. A quantitative estimation of the effect is possible only for asperities of very simple geometry [2].

Material displacement by plowing

When a hard asperity or an abrasive particle slides over the surface of a solid that is less hard, material is displaced by plowing and, as a result, grooves are formed. This process, which is typical for abrasive wear (see next section), requires energy and therefore contributes to friction.

© 2007, first edition, EPFL Press

Static and dynamic friction coefficients

Equation (10.7) provides a qualitative interpretation of the friction coefficient, which is the sum of three contributions: f_{ad} due to adhesion, f_{def} due to the elasto-plastic deformation of the asperities, and f_{pl} due to plowing:

$$f = f_{ad} + f_{def} + f_{pl} \tag{10.7}$$

The contribution of each mechanism to the friction coefficient depends on material properties, mechanical parameters and chemical conditions in the contact and cannot be predicted easily. Furthermore, their relative importance can change during the course of an experiment for several reasons: accumulation of wear particles forming a third body; change in rugosity due to plastic deformation of the asperities; strain hardening of the metal near the surface; increase of the contact temperature due to friction, etc.

The friction coefficient discussed above refers to steady state motion of the contact surfaces. The tangential force required to initiate sliding of a solid at rest is normally higher than that required to maintain the sliding motion. We therefore make a distinction between the ***static friction coefficient***, which refers to the minimum tangential force needed to initiate sliding of contacting surfaces at rest, and the ***dynamic friction coefficient***, which refers to the tangential force needed to maintain sliding at constant speed. The values listed in Table 10.3 refer to dynamic friction coefficients.

Contact temperature

Friction between two surfaces transforms mechanical energy into thermal energy, which leads to heating of the contact. When sliding starts, the contact temperature increases until it reaches a value where the heat production by friction equals the heat loss by thermal conduction.

Figure 10.7 shows a simple one-dimensional model of the contact between two materials sliding over another at constant speed. Material 1, which is assumed to dissipate most of the heat exhibits a linear temperature profile. The amount of thermal

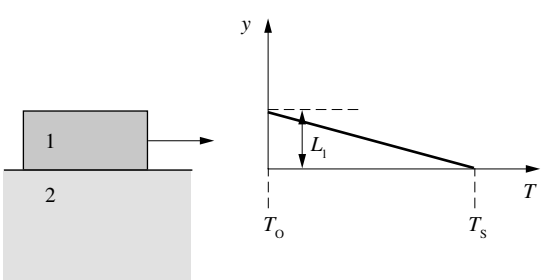


Figure 10.7 Average temperature profile in a sliding contact.

© 2007, first edition, EPFL Press

energy produced, Q_f (J m^{-2} s^{-1}), is proportional to the frictional force F, and to the sliding velocity v_{sl}. With (10.1), we obtain:

$$Q_f = \frac{f\,F_N\,v_{sl}}{A} \tag{10.8}$$

The heat flux Q_1 (J m^{-2} s^{-1}) in material 1 is equal to:

$$Q_1 = -k_1 \left.\frac{dT}{dy}\right|_{y=0} \approx k_1 \frac{T_s - T_o}{L_1} \tag{10.9}$$

where k_1 represents the thermal conductivity and y is the direction perpendicular to the contact surface. Furthermore, T_s and T_o are the temperatures, respectively, at the contact surface and at a characteristic distance L_1, the thermal conduction length. In Figure 10.6, L_1 is equal to the thickness of material 1.

The thermal flux Q_1 only represents a fraction θ_1 of the total heat dissipated Q_f, because material 2 also contributes. For steady state conditions we therefore write:

$$Q_1 = \theta_1\,Q_f \tag{10.10}$$

With the relation (10.9), the average temperature a the contact surface then becomes:

$$T_s = T_o + \frac{\theta_1\,L_1\,f\,F_N\,v_{sl}}{k_1\,A} \tag{10.11}$$

The surface temperature is proportional to the friction coefficient f, to the normal force F_N, and to the sliding velocity v_{sl}. The value of θ_1 depends on the ratio of the thermal conductivities of the contact materials, for identical materials, it is about 0.5.

In reality, the temperature profile is not necessarily linear and unidirectional, but depends on the geometry of the contact arrangement. The thermal conduction length L_1 is therefore not known with precision and the value of θ_1 may also be affected. In spite of these restrictions, the above model is useful because it shows how different variables affect the contact temperature.

In extreme cases, the temperature at the rubbing the surfaces can reach very high values, potentially leading to phase transitions in the metal adjacent to the surface or even to melting, especially if low-melting metals such as lead, tin or zinc participate in the contact.

Transient temperature at asperity contacts

Equation (10.11) reflects the energy balance over the entire contact area. The temperature T_s therefore corresponds to the average temperature at the contacting surfaces. And yet, because of their rugosity, the surfaces only touch at the asperities. During sliding, asperity contacts are continuously formed and broken and therefore two asperities remain in contact only for a short time. During this time, a significant amount of mechanical energy is transformed into heat in a small space. This leads to

© 2007, first edition, EPFL Press

temperature spikes at the contacting asperities, where for short times the transient temperature greatly exceeds the average surface temperature T_s. High peak temperatures can locally modify the properties of the material and accelerate chemical reactions.

By making certain simplifying assumptions, it is possible to estimate typical values of peak temperature [3]. However, the precision of such calculations is limited because the temperature spikes depend not only on sliding velocity and normal force, but also on parameters whose effects are difficult to quantify, such as surface topography, asperity deformation, oxidation kinetics and transient heat dissipation.

Seizure

From a practical point of view we make a distinction between "normal" conditions of friction and wear and "catastrophic" conditions that cause a drastic increase of the friction coefficient and lead to rapid destruction of the contact. The described phenomenon called ***seizure*** normally appears only after some time because the contact gradually deteriorates until adhesive junctions cover a significant fraction of the surfaces and the contact abrupt fails when the surfaces stick together.

The increase in contact temperature due to frictional heating generally favors the occurrence of seizure. Certain wear tests used in practice for the development of lubricants are based on the measurement of a critical normal force that causes seizure under increasingly severe wear conditions.

10.2.3 Friction in lubricated contacts

The presence of solid or liquid lubricants considerably reduces the friction coefficient, whose value then drops into a range typically between 0.02 and 0.1. Figure 10.8 gives a schematic representation of the three regimes of liquid lubrication:

- hydrodynamic lubrication;
- mixed lubrication;
- boundary lubrication.

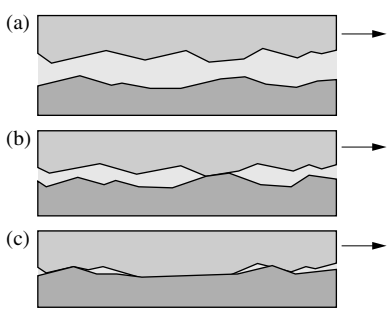


Figure 10.8 Lubrication regimes: (a) hydrodynamic lubrication; (b) mixed lubrication; and (c) boundary lubrication.

© 2007, first edition, EPFL Press

In the ***hydrodynamic lubrication*** regime the properties of the continuous lubricant film that separates the two surfaces, in particular its viscosity, determine the friction coefficient. The average thickness of the liquid film h is much larger than the average roughness R_a of the contact surfaces:

$$\frac{h}{R_a} >> 1 \tag{10.12}$$

Hydrodynamic lubrication is favored by a high sliding velocity, a weak normal force and a high viscosity of the lubricant. In machines, lubricated contacts normally operate in the hydrodynamic regime, minimizing thus friction and wear.

At higher normal force, or at lower lubricant viscosity and velocity, the ***mixed lubrication regime*** applies. The thickness of the liquid film, in this case, is comparable to the average surface roughness; some asperities can therefore touch one another.

$$\frac{h}{R_a} \approx 1 \tag{10.13}$$

Finally, in the ***boundary lubrication*** regime the average thickness of the liquid film is significantly smaller than the average surface roughness.

$$\frac{h}{R_a} << 1 \tag{10.14}$$

A large number of the asperity contacts are formed and the lubricant loses most of its efficiency. The boundary lubrication regime is favored by a high normal force, a low velocity and a low viscosity of the lubricant. Under these conditions, high friction often leads to failure of the contact by seizure.

Stribeck curve

The Stribeck diagram (Figure 10.9) presents the friction coefficient of a lubricated contact as a function of a parameter S, which has units of length.

$$S = \frac{v_{sl}\,\eta}{F_N} \tag{10.15}$$

Here, η designates the viscosity of the lubricant, v_{sl} is the sliding velocity, and F_N is the normal force. The positions of the different lubrication regimes on the Stribeck curve are indicated in Figure 10.9. For hydrodynamic lubrication, the friction coefficient can be described by the relation (10.16), where k_f is a proportionality factor.

$$f = k_f S^n \tag{10.16}$$

On a semi-logarithmic plot this corresponds to a straight line of slope n, where n typically has a value between 0.5 and 1. Deviations from linearity near the minimum

© 2007, first edition, EPFL Press

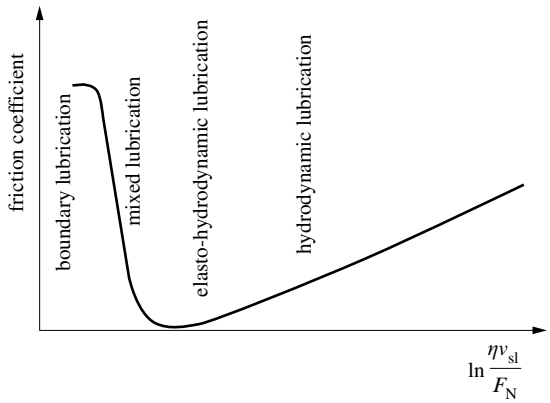


Figure 10.9 Stribek curve: Friction coefficient as a function of $\eta \upsilon_{sl}/F_N$, where η is the viscosity, υ_{sl} is the sliding velocity, and F_N is the normal force.

occur (*elasto-hydrodynamic lubrication*) if the contacting materials undergo elasto-plastic deformation or if the viscosity of the lubricant changes.

Surface chemical reactions of oil additives

Lubricating oils usually contain dissolved additives that reduce friction and wear. Anti-wear additives protect metal surfaces mostly in the mixed lubrication regime and to some extent also in the boundary lubrication regime. They react chemically with the surface, forming reaction products that prevent metal-to-metal contact at the asperities and thus facilitate sliding. Oil additives are typically metallo-organic or organic compounds that contain one or several polar functional groups, such as phosphorous esters. The tribochemical reaction between a phosphorous ester based additive and a steel surface yields primarily iron phosphate that stays on the surface.

In order to illustrate the effect of such additives, Figure 10.10 shows the variation of the friction coefficient as a function of the temperature of the lubricant, in the presence of an organic additive of the phosphorous ester variety. The increasing temperature gradually reduces the viscosity of the lubricant and the friction coefficient therefore decreases. At a temperature of about 80 °C, a transition occurs and the value of the friction coefficient abruptly increases from 0.09 to 0.12, due a change in surface chemistry. Auger electron spectroscopy analysis of the contact surfaces after completion of the experiments confirms that below the transition temperature, a thin layer of iron phosphate is present on the steel surface, visible on Figure 10.10 as an intense phosphorus signal. At the transition temperature, the intensity of the phosphorus signal decreases, and sulfur appears. Apparently, iron sulfides are now formed preferentially by a reaction of the solvent (that contains organic sulfides) with the metal surface. The iron sulfides are less efficient in reducing friction than the phosphates and as a consequence the friction coefficient abruptly increases. In such experiments, the transition temperature varies with the molecular structure of the

© 2007, first edition, EPFL Press

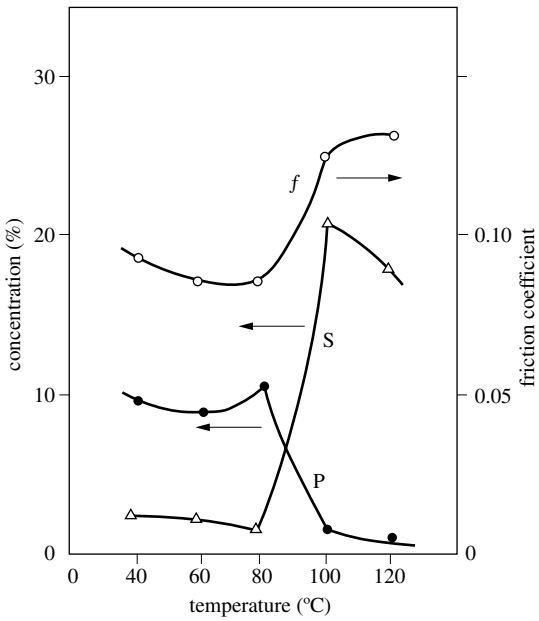


Figure 10.10 Effect of additives on friction coefficient of a lubricated steel surface, and sulfur and phosphorus surface concentration measured at a depth of 4.3 nm by Auger electron spectroscopy [4].

additive, which determines its thermal stability and its reactivity with the metal surface.

Friction in water lubricated systems

Water lubricated tribological systems are encountered in a variety of applications such as marine technology, food processing and biomedical implants. Water has a much lower viscosity than lubricating oils, which favors boundary lubrication conditions. Furthermore, in presence of ionic species an electrolyte is formed that facilitates corrosion. The value of the friction coefficients observed in water lubricated systems depends on the surface state of the metal, among other, and hence on the prevailing electrochemical conditions. This is illustrated by Figure 10.11 which shows the friction coefficient measured in a potentiodynamic experiment involving sliding of an aluminum oxide pin on a stainless steel plate in 0.5 M sulfuric acid. While sweeping the potential in the anodic direction, the current and the friction coefficient were determined. The results show that at the active-passive transition the friction coefficient decreases markedly due to the formation of the passive film. At higher potentials it increases again slightly, possibly because of increased formation of wear particles. The data of Figure 10.11 exemplify how a modification of the chemical state of a metal surface brought about by electrochemical polarization can modify its friction behavior.

© 2007, first edition, EPFL Press

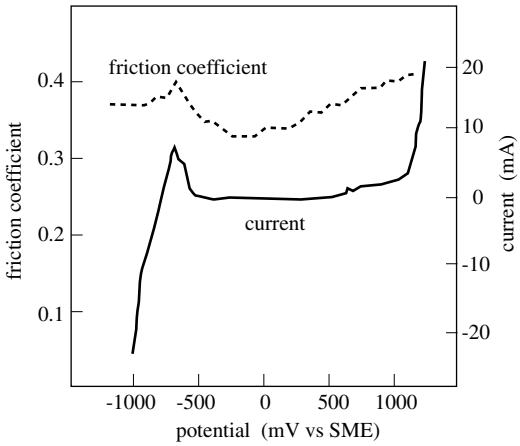


Figure 10.11 Variation with applied potential of the friction coefficient between an AISI 430 stainless steel and an aluminium oxide pin in 0.5 M sulfuric acid. Also shown is the anodic polarization curve of the steel [5].

10.3 WEAR AND CORROSION IN SLIDING CONTACTS

10.3.1 Experimental methods

The extrapolation of the results of accelerated wear tests to real operating conditions poses many problems and a multitude of test equipments have been developed for specific purposes. Among these, the following arrangements are most common (Figure 10.12):

- pin-on-disk;
- crossed-cylinder;
- pin-on-plate in reciprocating motion.

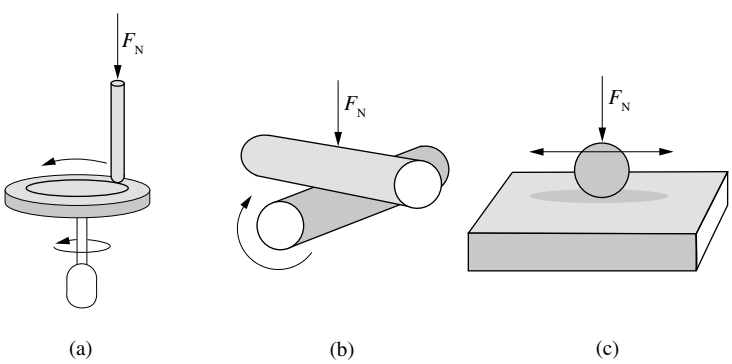


Figure 10.12 Experimental arrangements for tribological testing: (a) pin-on-disk; (b) crossed cylinders (Reichert); and (c) ball-on-plate.

© 2007, first edition, EPFL Press

Pin-on-disk devices consist of a rotating disk onto which a pin or ball is pressed at a constant normal force. Measurement of the normal force and of the angular momentum permits the determination the friction coefficient.

Crossed cylinder testing uses two perpendicular cylinders, one stationary, and the other in rotation. As in the previous case, the normal force and the angular momentum provide the information required to get the friction coefficient.

A reciprocating motion set-up involves the back-and-forth motion between a ball or a pin and a flat plate. The friction coefficient is obtained from the measurement of the normal and the tangential forces. This type of set-up can be easily adapted to the study of tribocorrosion behavior in electrolyte solutions. For this, the contact is immersed in the electrolyte that is contained in a cell equipped with a reference electrode and a counter electrode. The rubbing metal surface is branched as the working electrode in a potentiostatic circuit, which permits to perform rubbing experiments under potential control. The friction coefficient shown in Figure 10.11 was determined in this way by scanning the potential of the stainless steel in the positive direction.

Wear rate

One distinguishes the instantaneous and the average wear rate. The instantaneous wear rate during a test is obtained by recording the vertical displacement of the pin as a function of time, using suitable sensors. Another method consists of continuously monitoring the quantity of lost material using analytical or radiotracer methods. Most often though, one determines only an average wear rate over the time of an experiment. For this purpose, one measures the volume of worn material (*wear volume*) at the end of a test and divides it by the time of rubbing. Different quantities can be measured to obtain the wear volume:

- the mass loss of the sample;
- the volume of the wear trace;
- the total vertical pin displacement;
- the quantity of wear debris and dissolution products formed.

In order to get information on wear mechanisms, it is necessary to complement the determination of wear volume by a structural characterization of the worn surface. Friction and wear cause characteristic changes of surface topography and of the microstructure in a thin zone below the rubbing surface, which can give information on prevailing wear mechanisms. Additional information is obtained by an analysis of the size and composition of the wear particles. Electron microscopy (SEM, TEM) and surface analysis methods (AES, XPS, etc) are generally used for this purpose.

A major problem of wear studies is reproducibility. To get reproducible results, one needs first of all a mechanically well designed equipment that is sufficiently rigid to permit precise control of the mechanical variables. In addition, one needs to rigorously control the operational and environmental parameters, including humidity and temperature.

© 2007, first edition, EPFL Press

10.3.2 Wear mechanisms

The rate of wear v_w (m^3/m) is usually expressed as the derivative of the wear volume V_w with respect to the sliding distance l_{sl}:

$$v_w = \frac{dV_w}{dl_{sl}} \tag{10.17}$$

According to ***Archard's law***, found empirically, the wear volume is proportional to the normal force F_N and to the sliding distance l_{sl} and inversely proportional to the hardness H of the softer of the two materials in contact:

$$V_w = K_w \frac{F_N l_{sl}}{H} \tag{10.18}$$

The dimensionless proportionality factor K_w is called the ***wear coefficient***. With equation (10.17), we get for the rate of wear:

$$v_w = K_w \frac{F_N}{H} \tag{10.19}$$

The physical interpretation of the wear coefficient K_w depends on the dominating wear mechanism. Four types are commonly distinguished in the literature:

- adhesive wear: formation of adhesive junctions and tearing off of metal;
- abrasive wear: plowing by surface asperities or abrasive particles;
- delamination wear: growth of sub-surface cracks leading to fracture by fatigue;
- oxidative wear: growth and detachment of oxide films.

Adhesive wear

Adhesive wear can be explained by a model first proposed by Archard [6]. The two surfaces in relative motion only touch at the asperities. When the normal force F_N is applied, the contact zones undergo plastic deformation and micro-welds, referred to as *adhesive junctions*, are formed. This is the same mechanism as that governing adhesive friction, but here we are interested not primarily in energy dissipation, but in the rate at which material is torn off from the adhesive junctions. If n_j is the number of adhesive junctions, the contact area is given by:

$$A_r = n_j \pi r_j^2 \tag{10.20}$$

where r_j represents the average radius of the junctions, assumed to be circular. The relation (10.4) allows us to express A_r in terms of the hardness H and the normal force F_N. The number of junctions thus is:

$$n_j = \frac{F_N}{H \pi r_j^2} \tag{10.21}$$

© 2007, first edition, EPFL Press

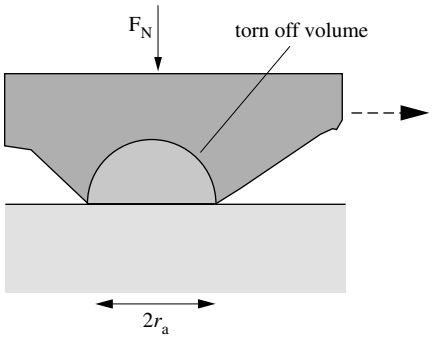


Figure 10.13 Adhesive wear model.

The lifetime t_j of a junction amounts to:

$$t_j = \frac{2\,r_j}{v_{sl}} \tag{10.22}$$

where v_{sl} designates the sliding velocity. The number of junctions formed over a given sliding length, N_j, is equal to:

$$N_j = \frac{n_j}{t_j\,v_{sl}} = \frac{n_j}{2\,r_j} \tag{10.23}$$

Certain junctions can break without resulting in material removal. The parameter K_{wA} designates the probability that a junction that breaks will lead to formation of a wear particle. In order to simplify, we assume that the volume that breaks off from an asperity is equivalent to a hemisphere of radius r_j (Figure 10.13). The rate of wear v_w is then obtained by multiplying this volume by the number of junctions formed and by the probability K_{wA}:

$$v_w = K_{wA}\,N_j\,\frac{4\pi\,r_j^3}{6} \tag{10.24}$$

With (10.24), (10.20) and (10.4) we get for the wear rate:

$$v_w = \frac{dV_w}{dt} = \frac{K_{wA}}{3}\,\frac{F_N}{H} \tag{10.25}$$

and after integration:

$$V_w = \frac{K_{wA}}{3}\,\frac{F_N\,l_{sl}}{H} \tag{10.26}$$

This corresponds to Archard's law (10.18), with $K_w = K_{wA}/3$. In the literature, the term "wear coefficient" is sometimes used for K_{wA}. However, in order to avoid

© 2007, first edition, EPFL Press

confusion, we will call K_{wA} *Archard coefficient*. It indicates the probability that an adhesive junction will lead to the formation of a wear particle. The Archard coefficient has a precise physical meaning based on a specific theoretical model, in contrast to the wear coefficient, which is an experimentally determined proportionality factor. The model for adhesive wear presented here assumes a simple geometry for the wear particles that break off. Models using more realistic geometries have been proposed, but the results are not fundamentally different.

The described model of adhesive wear does not take into account oxide films or adsorbed species that are generally present at metal surfaces. They are expected to affect the bond energy at asperity junctions or even to inhibit formation of metal to metal junctions all together. The rate of wear should therefore depend on the chemical environment. This is confirmed by the results presented in Table 10.14: rubbing of copper surfaces produces wear particles of different size, depending on the environment. The diameter of the particles in this case is taken as a qualitative indication of the wear volume, assuming that a large diameter indicates a large wear volume. The results indicate that the presence of oxygen or humid air, which favor surface oxidation, reduces the wear rate of a copper-copper contact. A lubricant reduces it even further.

Table 10.14 Average diameter of wear particles generated by rubbing of two copper surfaces exposed to different environments [8].

Environment	Average diameter in μm
nitrogen	480
helium	380
carbon dioxide	300
dry air	224
oxygen	201
humid air	144
liquid lubricants	8 – 12

Abrasive wear

Abrasive wear involves the displacement of material by plowing and severe plastic deformation. Eventually junks of displaced material break off and form a wear particle.

Abrasive wear occurs whenever a rough hard material slides over a softer material. It is easily identified by the presence of grooves and scratches on a worn surface. Motors, gears and ball bearings can undergo abrasive wear in the presence of wear particles that have been hardened by cold deformation or oxidation. Abrasive papers or powders are widely used in surface finishing to produce a desired surface topography by controlled abrasive wear. A typical example is mechanical polishing with diamond powders.

© 2007, first edition, EPFL Press

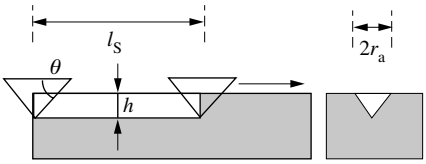


Figure 10.15 Abrasive wear model.

Figure 10.15 presents a simple model for abrasive wear. A conical asperity made of hard material moves across a soft surface. By action of the normal force F_N, the asperity penetrates into the surface to a depth h,

$$h = r_j \tan\theta \tag{10.28}$$

where r_j represents the contact radius and θ corresponds to the angle indicated in the Figure. After the asperity has moved through a distance l_{sl}, it has displaced the volume $V_w = l_{sl}\, h\, r_j$. By setting

$$\pi r_j^2 \approx \frac{F_N}{H} \tag{10.28}$$

we obtain:

$$V_w = \frac{\tan\theta}{\pi}\,\frac{F_N\, l_{sl}}{H} \tag{10.29}$$

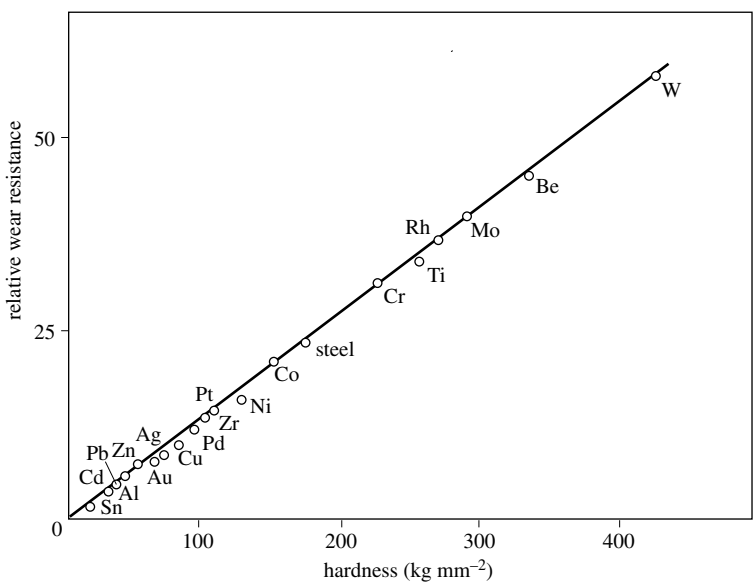


Figure 10.16 Effect of hardness on wear resistance of pure metals [8].

© 2007, first edition, EPFL Press

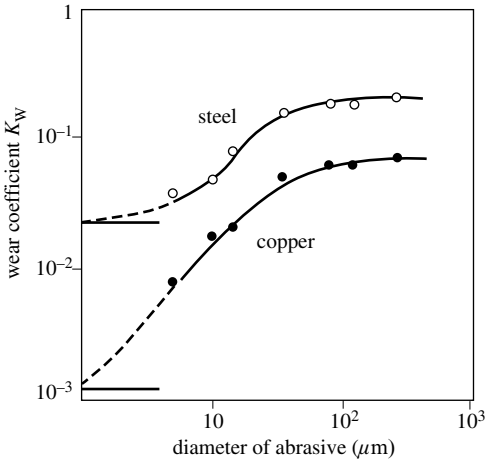


Figure 10.17 Effect of diameter of abrasives on the wear coefficient [8].

This last equation corresponds to the relation (10.18) with

$$K_w = \frac{\tan\theta}{\pi} \tag{10.30}$$

The wear coefficient, according to this model, therefore depends only on the angle θ that characterizes the geometry of attack by the abrasive. In reality, θ must be viewed as an average over many different angles of attack, and its value is usually not known.

The model predicts that the rate of abrasive wear is inversely proportional to the hardness of the material under attack. Experimental results obtained with pure metals seem to agree with this prediction (Figure 10.16). On the other hand, for metals such as steel, whose hardness varies with the microstructure, the relation (10.29) does not hold well. The hardness of the abrasive does not appear in equation (10.29). Experience indicates that the hardness of the abrasive can be neglected if it is at least 1.3 times that of the abraded metal. Neither does the size of the abrasive particles critically affect the rate of wear if it exceeds a few tens of micrometer. Figure 10.17 illustrates the influence of the particle size of an abrasive paper on the wear coefficient, measured on carbon steel and copper. The value of K_w increases up to a particle size of about 50 μm, and then stabilizes. Under certain conditions, the abrasive particles themselves may undergo wear or break up during a test. The wear coefficient then decreases gradually with time as the particle size decreases and the proportionality between the wear volume and sliding distance predicted by equation (10.29) is no longer valid.

Delamination wear

Under certain conditions, fatigue cracks form in a plane parallel to the contact surface at a depth of several micrometer. When they reach a critical size, local fracture

© 2007, first edition, EPFL Press

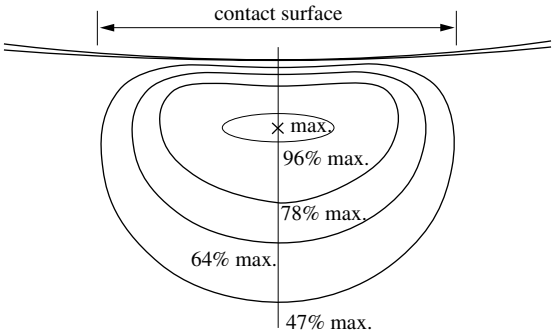


Figure 10.18 Stress distribution in sphere-to-plate Hertzian contact [9].

produces a wear particle. Delamination wear therefore is also known as ***fatigue wear***. The reason that cracks form at some distance below the surface is that the shear stress there is highest. The stress distribution for a Hertzian contact illustrates this fact (Figure 10.18). The Hertzian model considers elastic stresses resulting from the action of a rigid sphere on a plane under static conditions. When the two surfaces slide on each other, the stress distribution becomes more complicated due to the presence of shear stress; but for most usual conditions, the maximum is still located below the surface.

Different microscopic mechanisms can contribute to fatigue crack formation. One plausible hypothesis postulates that cracks originate from cavities that are present in the vicinity of non-metallic inclusions [10]. Figure 10.19 shows this model schematically. A rough, hard surface slides over a metal containing rigid non-metallic inclusions. The asperities of the rough surface are separated by the average distance λ. Each contact of an asperity with the metal causes an incremental shear deformation γ_o. The maximum shear is observed in a plane parallel to the surface and located at the depth h, assumed to be of the same order of magnitude (several micrometer) as

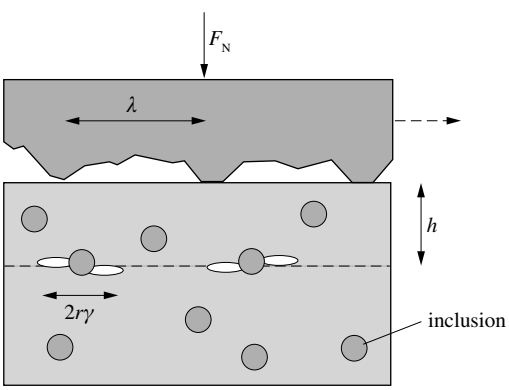


Figure 10.19 Delamination wear model: cracks develop at inclusions situated at a depth h from the surface [10].

© 2007, first edition, EPFL Press

the radius r_j of the asperities. Cracks first appear at the cavities located in this plane, and their size increases with each stress cycle. After a critical number of cycles (each corresponding to a microscopic contact) , n_{crit} , the cracks reach a critical size that leads to fracture and formation of a wear particle. A volume of material equal to $A\,h$ detaches, where A represents the geometrical surface. The critical number of cycles is reached after a sliding distance $n_{crit}\,\lambda$. The wear rate therefore is equal to:

$$v_w = \frac{A\,h}{n_{crit}\,\lambda} \tag{10.31}$$

By making the approximations $A/A_r \approx \lambda/r_j$ and $h \approx r_j$ where A_r designates the real contact area, we find for the wear rate:

$$v_w = \frac{A_r}{n_{crit}} \tag{10.32}$$

Setting as before $A_r = F_N/H$ and integrating over the sliding distance l_{sl} yields the wear volume.

$$V_w = \frac{1}{n_{crit}}\,\frac{F_N\,l_{sl}}{H} \tag{10.33}$$

This equation corresponds to Archard's law with $K_w = 1/n_{crit}$. The wear coefficient in this case represents the inverse of the critical number of asperity contacts that lead to fracture by fatigue. At each passage of an asperity the subsurface crack grows by an increment until it reaches a critical size at which fracture occurs. The quantity n_{crit} can be associated to the number of cycles leading to fracture in a common fatigue test. Many parameters affect the value of n_j in a wear test and it is not possible to calculate it from first principles. For the model structure shown in Figure 10.19, the critical number of cycles can be brought in relation with the number of inclusions or precipitates that act as crack initiation sites [10].

Oxidative wear

Oxidative wear takes place in the presence of oxygen or other oxidizing gases at high temperature. High temperatures in the contact zone may be the result of an elevated temperature outside or of excessive frictional heating. High temperatures facilitate the growth of relatively thick oxide films by solid state diffusion (Chap. 9). During rubbing, when the film thickness reaches a critical value, the oxide breaks off under the effect of shear forces. A new oxide layer then forms on the exposed metal surface, and the cycle continues.

Figure 10.20 schematically represents the described behavior. A hard asperity moves with respect to a rough surface covered with an oxide layer that then detaches at the contact point. This simple model allows us to estimate the rate of oxidative wear [11]. For this we assume that the oxide growth follows a parabolic rate law:

© 2007, first edition, EPFL Press

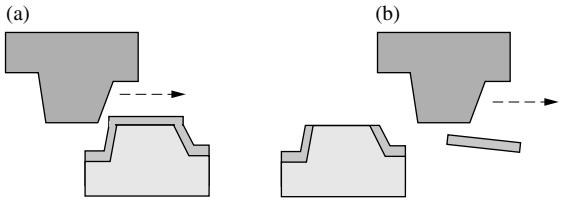


Figure 10.20 Oxidation wear model: (a) An oxide layer grows to a critical thickness and (b) is removed from the surface by shear forces.

$$L^2 = k_L t \tag{10.34}$$

where L represents the oxide thickness, t is the time and $k_L = k_p/\rho_{ox}{}^2$ is a parabolic rate constant (cf. equation 9.49). The layer reaches a critical thickness L_{crit} after a time t_{crit}:

$$t_{crit} = \frac{L_{crit}^2}{k_L} \tag{10.35}$$

At time t_{crit}, which corresponds to the sliding length $l_{crit} = t_{crit}\, v_{sl}$, a volume of oxide $A_r\, L_{crit}$ breaks off from the surface. Here v_{sl} designates the sliding velocity and A_r the real contact area. If the specific volume of the oxide is approximately equal to that of the metal (if not, it would be simple to introduce a correction factor), the wear rate (in m^3 per m) is equal to:

$$v_w = \frac{dV_w}{dl_{sl}} \approx \frac{A_r\, L_{crit}}{v_{sl}\, t_{crit}} \tag{10.36}$$

After replacing t_{crit} by (10.35) and setting: $A_r = F_N/H$, the integration of (10.36) gives:

$$V_w = \frac{k_L}{v_{sl}\, L_{crit}} \frac{F_N\, l_{sl}}{H} \tag{10.37}$$

This expression agrees with Archard's law, with

$$K_w = \frac{k_L}{v_{sl}\, L_{crit}} \tag{10.38}$$

In this model, the wear coefficient depends on the sliding rate v_{sl} and on the oxidation rate constant k_L. The latter is a function of the contact temperature and follows Arrhenius' law, $k_L = k_{L,0} \exp(-\Delta G^{\#}/RT)$, where $\Delta G^{\#}$ is the activation energy.

In reality, the wear behavior is more complicated than predicted by this model. Firstly, the oxides are unlikely to break off from the metal in one piece once they reach a critical thickness. Rather, oxide fragments of variable shape and size will be

© 2007, first edition, EPFL Press

produced from which the film thickness can not be readily estimated. In addition, the released oxide particles may form a third body and depending on their hardness, cause abrasive wear or act as a lubricant.

Comparison of wear mechanisms

The four wear mechanisms described in this section all lead to Archard's law, but the physical interpretation of the wear coefficient differs. In the adhesive wear model the wear coefficient expresses the probability that an adhesive junction leads to formation of a wear particle. In the abrasive wear model the wear coefficient depends only on the geometry of the abrasive. The wear coefficient in delamination wear characterizes the critical number of cycles leading to fatigue fracture of microscopic subsurface cracks. Finally, the wear coefficient in oxidative wear includes the growth constant and the critical thickness of a surface oxide film.

Which wear mechanism dominates in a given situation is not always obvious and often two or more wear mechanisms may act jointly. Because all models predict the same dependency on hardness, normal force and sliding velocity, other aspects must be studied to distinguish wear mechanisms. Useful information on wear mechanisms is obtained from microstructural observation of the worn surfaces and of wear debris. Finally, we note that three of the above wear mechanisms involve only mechanical interactions, while oxidative wear is in fact a tribocorrosion phenomenon, combining mechanical wear and high temperature oxidation.

Wear maps

Wear maps present in graphical form the predominant wear mechanism as a function of selected system variables, usually written in dimensionless form in order to minimize their number. Figure 10.21 presents a wear map, for pin-on-disk tests of non-lubricated steel to steel contacts. The abscissa shows the dimensionless sliding velocity v_{sl}^*, and the ordinate gives the dimensionless normal force F_N^*. These two parameters correspond, respectively, to the two following definitions:

$$v_{sl}^* = \frac{v_{sl}\, l_{sl}}{\alpha_T} \tag{10.39}$$

$$F_N^* = \frac{F_N}{A\,H} \tag{10.40}$$

The dimensionless sliding velocity contains the thermal diffusivity of the metal, α_T, which influences the contact temperature resulting from frictional heating.

In Figure 10.21, the thick lines define six domains of wear: (I) mild wear, (II) wear dominated by the elasto-plastic behavior of the metal (delamination, adhesion, abrasion), (III) oxidation wear, (IV) oxidation wear with oxide melting, (V) wear by surface melting of the metal, (VI) seizure. The thin lines crossing the domains represent constant, dimensionless wear rates. The dimensionless wear rate v_w^* is equal

© 2007, first edition, EPFL Press

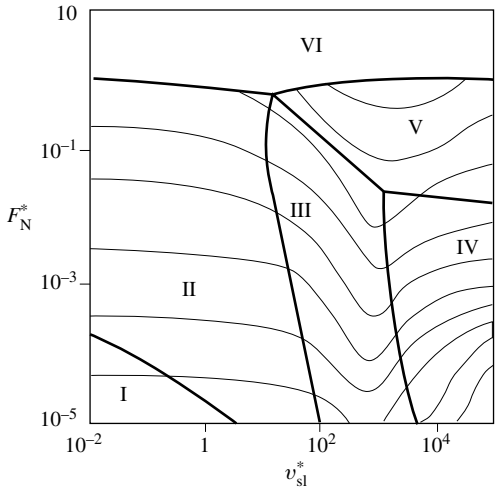


Figure 10.21 Wear map of steel for pin-on-disk experiment showing the different domains of wear [10].

to the wear rate v_w en m^3/m, divided by the geometrical surface area of the contact, A:

$$v_w^* = \frac{v_w}{A} \tag{10.41}$$

The wear map of Figure 10.21 corresponds to a particular case, that is, a non-lubricated steel to steel contact in a pin-on-disk apparatus. This situation is of little practical interest, because most metal to metal contacts are lubricated. Furthermore, much of the parameter space covered in the map refers to extreme conditions, far from typical engineering situations. The wear map shown is nevertheless interesting because it provides an overview of how a combination of different variables can affect the wear mechanism and the wear rate.

10.3.3 Tribocorrosion behavior of passive metals

Passive film damage due to rubbing

Metals and alloys used in engineering are usually in their passive state. When a passive metal in presence of a corrosive electrolyte rubs against a solid, both corrosion and wear occur, but because the mechanical and electrochemical mechanisms interact with each other, the overall rate of material loss is not just the sum of material lost by corrosion and by wear taken individually.

Let us assume that a hard pin slides over a passive metal surface immersed in an electrolyte**.** Depending on the severity of the wear conditions rubbing will lead to thinning of the passive film in the contact zone or to its complete removal at contacting asperities, exposing locally the film free metal surface to the electrolyte. Rapid

© 2007, first edition, EPFL Press

oxidation will restore the passive film once the asperity contact is broken. Film growth may be preceded or accompanied by active metal dissolution. Three processes then contribute to material removal from the rubbing metal: growth of the passive film, metal dissolution into the electrolyte and mechanical wear leading to detachment of particles.

In a freely corroding system exposed to oxygen or some other oxidizing agent, the electrons liberated by the partial anodic reaction will be consumed by the reduction of the oxidizing agent. Because rubbing locally damages the passive film the two partial reactions do not take place uniformly over the surface. The anodic reaction will be located predominantly in the wear scar and the cathodic reactions on the outside surface. Thus rubbing leads to the formation of a corrosion cell between the wear scar and the untouched passive surface, the wear scar forming the anode. Indeed, passive film damage lowers the local corrosion potential in the wear scar and a galvanic current flows between the wear scar and the undamaged surface.

The corrosion potential measured by means of a reference electrode placed in the vicinity of the contact changes in the negative direction when rubbing starts (Figure 10.22). When rubbing stops it returns to its original value. The potential measured during rubbing is not the actual corrosion potential of the wear scar but includes the ohmic potential drop between the point of measurement and the wear scar, which depends on the galvanic current flow between anodic and cathodic areas. The data of Figure 10.22 nevertheless clearly demonstrate that the measured corrosion potential during rubbing is much lower than that of the intact passive surface, confirming that rubbing locally depassivates the surface.

In the laboratory we can replace the oxidizing agent by an external current source. For this the rubbing metal surface is made the working electrode in an electrochemical cell that contains a counter and a reference electrode. In the absence of other reactions, the anodic current in a potentiostatic experiment is then a direct measure of the rate of

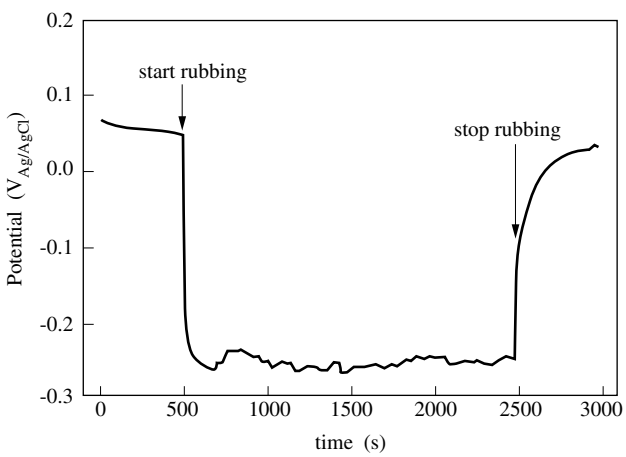


Figure 10.22 Effect of rubbing on corrosion potential. AISI 304 stainless steel plate sliding against aluminium oxide ball in 0.5 M NaCl at open circuit [12].

© 2007, first edition, EPFL Press

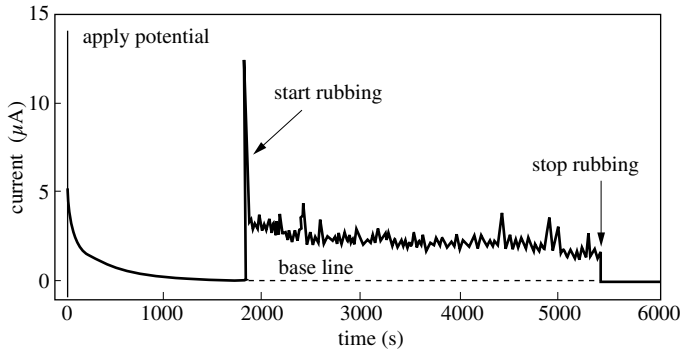


Figure 10.23 Effect of rubbing on the current measured on Ti6Al4V alloy in 0.9 w% NaCl, anodically polarized in the passive potential region. Counter piece: aluminium oxide ball. Applied potential: 0.5 $V_{Ag/AgCl}$ [13].

metal oxidation. Electrochemical experiments thus permit us to measure in situ and in real time the rate of wear accelerated corrosion as a function of potential.

As an illustration, Figure 10.23 shows the current measured in a sliding contact between a titanium alloy plate and a polished aluminum oxide sphere in a 0.9 w% NaCl solution at a constant potential of 0.5V versus a silver/silver chloride reference electrode (passive potential region). When the titanium alloy is being immersed into the solution and the anodic potential is applied, one observes an anodic current transient resulting from passive film growth. As the film thickens the growth current decreases exponentially and eventually reaches a negligibly low value (Chapter 6). With the onset of rubbing, the anodic current increases abruptly because wear damages the passive film in the contact zone. When rubbing stops, the current drops again to a very small value as passivity is restored. The difference between the current measured during rubbing and the base line current yields the rate of wear accelerated corrosion, using Faraday's law.

Modeling wear accelerated corrosion

Let us consider first the following situation. A metal pin slides over a ceramic plate at a constant velocity in an electrolyte solution. The metal is polarized anodically in the passive potential region using a potentiostat. Rubbing causes wear of the passive oxide film by abrasion. Under steady state conditions the rate of film growth at the imposed potential is equal to the rate of film thinning due to wear:

$$\frac{dL}{dt} = 0 = k_F i_{gr} - K_w \frac{F_N}{H} v_{sl} \tag{10.42}$$

In this expression i_{gr} is the film growth current and $k_F = M/\rho nF$ is a constant taking into account Faradays law (c.f. Section 6.2.2). The second term represents the wear rate of the film in m/s. It is equal to the wear rate per unit length, v_w, given by equation (10.19) multiplied by the sliding velocity v_{sl}.

© 2007, first edition, EPFL Press

From (10.42) we obtain an expression for the anodic current that serves for film repair:

$$i_{gr} = K' \frac{F_N}{H} v_{sl} \tag{10.43}$$

with $K' = K_w/k_F$. The model predicts that the anodic current measured in an electrochemical tribocorrosion experiment at constant potential is proportional to the sliding velocity and to the ratio of normal force over hardness.

In developing equation (10.43) we assumed that wear causes film thinning, but not film removal. Under more severe wear conditions it is likely that the passive film is entirely destroyed at the asperity contacts. In the following, we shall consider a model where complete depassivation occurs at asperity contacts. As the asperity junctions are broken, the exposed surface areas repassivate again. We assume that a charge Q_p (C/m^2) per unit area is required to repassivate the metal and we call Q_p the ***repassivation charge.*** The anodic current is then given by

$$I_a = R_{dep}\ Q_p \tag{10.44}$$

where R_{dep} (m^2/s) is the rate of generation of depassivated surface due to rubbing. If the number of asperity junctions is n_j, their average area is A_j and their life time is t_j we have:

$$R_{dep} = k' \frac{n_j A_j}{t_j} \tag{10.45}$$

The proportionality factor k′ expresses the probability that a given asperity contact actually leads to local removal of the passive film. From (10.21), and assuming circular asperity contacts, we have $n_j = F_N/H\ \pi\ r_j^2$, and from (10.22) we find $t_j = 2r_j/v_{sl}$. This yields for the rate of generation of depassivated surface:

$$R_{dep} = \frac{k'}{2r_j}\left(\frac{F_N}{H}\right)v_{sl} \tag{10.46}$$

We can replace r_j in this formula by setting $r_j = A_j^{1/2}/\pi^{1/2} = (n_jA_j)^{1/2}/(\pi\, n_j)^{1/2}$. With the identity $n_jA_j = F_N/H$ we thus obtain:

$$R_{dep} = K_n\left(\frac{F_N}{H}\right)^{1/2} v_{sl} \tag{10.47}$$

with $K_n = (1/2)\ k'\ \pi^{1/2}\ n_j^{1/2}$. According to this relation, which supposes a constant number of asperity junctions, the rate of generation of depassivated area is proportional to the square root of F_N/H. A comparison of equations (10.46) and (10.47) shows that assuming the radius of the asperity junctions to be constant leads to a different result than assuming their number to be constant. In one case the ratio F_N/H enters into the final equation with an exponent of one, in the other case with an exponent of $\frac{1}{2}$. More generally, we can therefore write:

© 2007, first edition, EPFL Press

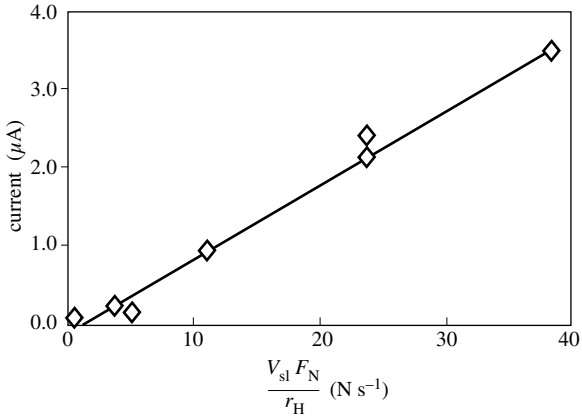


Figure 10.24 Average current measured during rubbing of Ti6Al4V alloy against an aluminium oxide ball plotted against the parameter $v_{sl}F_N/r_H$. Electrolyte: 0.9 w% NaCl. Applied potential: 0.5 V vs. Ag/AgCl [13].

$$I_a = R_{dep}\, Q_p = K_p \left(\frac{F_N}{H}\right)^b v_{sl} \tag{10.48}$$

where the value of the exponent b is $0.5 \le b \le 1$, depending on the particular conditions of the experiment and the surface topography [5]. The proportionality factor K_p can be associated with the wear coefficient.

In Figure 10.24 anodic current measured during rubbing of a titanium alloy against a polished aluminum oxide sphere is plotted against the quantity $v_{sl}F_N /r_H$, where r_H is the Hertzian radius of the impression made by the sphere. The data points, which were obtained at different applied normal force, sliding velocity and different stiffness of the equipment fall on a straight line. They are in good agreement with equation (10.46) provided that r_j is replaced by r_H. For the particular contact arrangement, involving a hard smooth surface sliding on a rough ductile material this assumption is reasonable.

Third body effects

The discussed models give a satisfactory interpretation of how mechanical and materials variables affect the rate of wear accelerated corrosion in a sliding contact, but they do not explicitly take into account the possibility that wear particles formed in the process may stay in the contact zone for some time before being ejected. The particles in the contact zone form a third body that changes the prevailing mechanical conditions.

To better understand the behavior of third body particles we consider the schematic picture shown in Figure 10.25. A metallic wear particle is formed by fatigue wear, one of the classical wear mechanisms. The metallic particle may either be

© 2007, first edition, EPFL Press

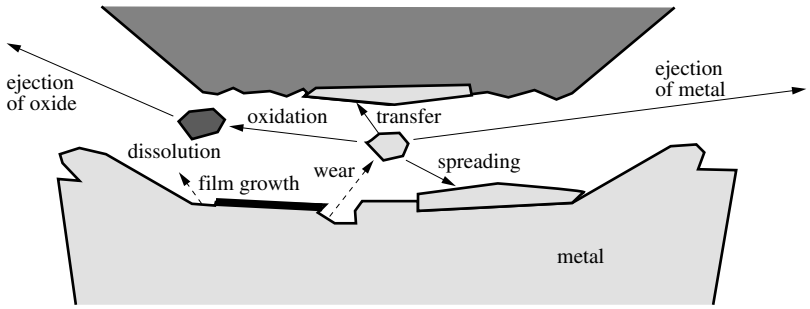


Figure 10.25 Schematic of third body behavior and material transfer processes in tribocorrosion when a hard pin slides over a metal.

directly ejected from the contact or it may first undergo a variety of chemical or physical transformations in the contact zone. Specifically, it may oxidize to an oxide particle that is then ejected or, alternatively, that dissolves in the electrolyte forming dissolved ions. Metallic third body particles may also undergo severe plastic deformation and be spread out on the rubbing metal before being ejected or they may stick to the counter piece forming a *transfer layer* that consists of either metal or oxide.

The described behavior of wear particles complicates the interpretation of electrochemical tribocorrosion experiments. In principle, the measurement of anodic current or charge permits one to identify the contribution of corrosion to the overall material loss from the rubbing metal surface. However, the electrochemical oxidation of metallic third body particles that are in electrical contact with the rubbing surface also contributes to the anodic current, although these particles were formed by mechanical wear [14]. To distinguish between electrochemical and mechanical metal removal mechanisms in tribocorrosion one therefore must take into account the behavior of third body particles.

A further complication arises if cathodic reactions take place in the contact zone or in its vicinity because in that case the actual oxidation rate in the contact zone exceeds the measured anodic current. The measured current $I_{measured}$ is the sum of the partial anodic current I_a, which represents the oxidation rate in the contact zone, and the (negative) partial cathodic current I_c: $I_{measured} = I_a + I_c$. Electrochemical methods are powerful tools for the study of tribocorrosion mechanisms in sliding contacts, but they must be complemented by structural and chemical investigations of wear particles and worn surfaces.

10.4 WEAR AND CORROSION DUE TO IMPACTS

10.4.1 Flow accelerated corrosion

Erosion corrosion is the combined action of erosive wear by impacting solids or liquids and of corrosion. In practice, the term erosion corrosion is often used broadly

© 2007, first edition, EPFL Press

to designate accelerated corrosion in turbulent flow systems, independently of whether solid particles are present or not. Corrosion in turbulent flow systems not containing suspended solids shall be referred to here as ***flow accelerated corrosion***.

Critical flow velocity

Certain metals corrode at an abnormally high rate when the flow velocity in the turbulent flow regime exceeds a critical value. The phenomenon is well known for copper and its alloys used in domestic hot water systems and in heat exchanger tubes. Table 10.26 indicates empirical values for the critical flow velocity of seawater in condenser pipes made of different alloys. Similar values apply for fresh water.

Table 10.26 Critical velocities for seawater in condenser pipes [15].

Metal	Critical flow velocity (m/s)
copper	0.9
Si-bronze	0.9
Admiralty brass	1.5
Cu-10Ni	3.0
Cu-30Ni	3.7
type 316 stainless steel	no limit observed
titanium	no limit observed

The occurrence of a critical flow velocity can be explained in two ways: by the action of shear stress on corrosion product films covering the pipe walls or by mass transport considerations. We first consider the role of shear stress.

The flow of a fluid causes a shear stress τ_0 at the wall:

$$\tau_0 = \eta \left. \frac{\mathrm{d}u_x}{\mathrm{d}y} \right|_{y=0} \tag{10.49}$$

Here, μ represents the viscosity, u_x the flow velocity parallel to the wall, and y is the direction normal to the wall. For a Newtonian fluid (constant viscosity) the shear stress is given by:

$$\tau_0 = \frac{1}{2} \rho u^2 f' \tag{10.50}$$

where ρ indicates the density of the fluid and v is its average speed in the direction of the flow. The friction coefficient f', dimensionless, represents the ratio between the frictional force F acting on the surface A and the kinetic energy of the fluid per unit volume, $E_{\mathrm{kin}} = (1/2)\, \rho\, v^2$:

$$f' = \frac{F / A}{E_{\mathrm{kin}}} \tag{10.51}$$

© 2007, first edition, EPFL Press

The value of f' depends on the flow velocity of the fluid, on the wall rugosity, and on the geometry of the installation.

Using equation (10.50) we can relate the critical flow velocity for erosion corrosion, v_{crit} to a ***critical shear stress*** τ_{crit}:

$$v_{crit} = \left(\frac{2\,\tau_{crit}}{\rho\, f'} \right)^{1/2} \tag{10.52}$$

If the critical shear stress that acts on the layer of corrosion products present at a surface is exceeded, mechanical film damage leads to a strong increase in the rate of corrosion. According to this view, stainless steel and titanium are not susceptible to flow accelerated corrosion because the thin passive oxide films formed on these metals are more resistant to shear stresses than the corrosion product layers found on copper and its alloys [16].

Unfortunately, no experimental proof has been given so far that flow accelerated corrosion is the result of exceeding a critical shear stress that acts on protecting surface films of corrosion products. Rather, a number of studies suggest that flow accelerated corrosion can be explained simply based on mass transport considerations [17].

For turbulent flow in pipes the mass transport coefficient and the friction coefficient are related by the ***Chilton-Colburn analogy***:

$$j_H = \frac{k_L}{u} Sc^{1/3} = \frac{f'}{2} \tag{10.53}$$

Here j_H is the Chilton-Colburn factor for mass transport, k_L is the mass transport coefficient, u the linear flow velocity, $Sc = \nu/D$ is the Schmidt number and f' is the friction coefficient. As a consequence, a higher shear stress also means a higher mass transport rate and vice versa.

In neutral media, the rate of corrosion is often partially or entirely governed by mass transport; either by the transport of an oxidizing agent, typically oxygen, towards the surface or by transport of a ionic reaction product away from the surface (Chapter 4). In the latter case the corrosion rate depends on the solubility of the dissolved ions, which by precipitation can form a salt film on the surface. The rate of mass transport under forced convection conditions, and hence the rate of corrosion, can be predicted by dimensionless relations of the type $Sh = f(Re,Sc)$, examples of which are given in Table 4.27. Complications arise, if the surface roughness increases during corrosion or if the corrosion rate is under mixed control, for example when a porous layer of corrosion products grows at the surface.

The data of Table 10.26 can now be interpreted in terms of mass transport. Above a critical flow velocity, which depends on pipe diameter, the rate of mass transport becomes sufficiently fast to carry away dissolved corrosion products without forming a salt film. In other words, their surface concentration remains below saturation. Under these conditions scales that are formed by precipitation of corrosion products at lower flow velocity can not exist and the corrosion rate therefore will be higher. An example in case is *carbon dioxide corrosion of carbon steel* that occurs in fluids containing

© 2007, first edition, EPFL Press

dissolved CO_2, typically found in the oil industry [18]. Protective scales of $FeCO_3$ precipitate if the solubility limit of Fe^{2+} and CO_3^{2-} ions at the surface is exceeded. This is favored by a low mass transport rate. On the other hand, at high flow rates accelerated corrosion occurs because protecting scales of $FeCO_3$ no longer form. The mass transport rate is now sufficient to evacuate ionic corrosion products without reaching the saturation limit at the surface.

High flow rates also increase the rate of transport of dissolved oxygen towards the pipe wall. For non-passive metals, such as copper, this means a higher corrosion rate at higher flow velocity. On the other hand, it is difficult to explain by this mechanism the observation of a critical flow velocity, as the corrosion rate should increase steadily with increasing flow velocity. As will be shown below, it is likely that several of the mentioned phenomena act together in flow accelerated corrosion of copper and its alloys. Among the metals listed in Table 10.26, stainless steel and titanium do not exhibit flow accelerated corrosion. In seawater these metals are passive and their corrosion rate is not limited by mass transport. On the contrary, accelerating the transport of oxygen helps to maintain passivity (Chapter 6).

In engineering systems local differences in mass transport rate are common. Examples are pipe bends, flow downstream from a expansion or complex flow fields in a centrifugal pump. Differences in mass transport rate can lead to local differences in the corrosion rate. Figure 10.27 shows mass transport rates measured downstream from a expansion in a straight pipe by determining the limiting current for ferricyanide

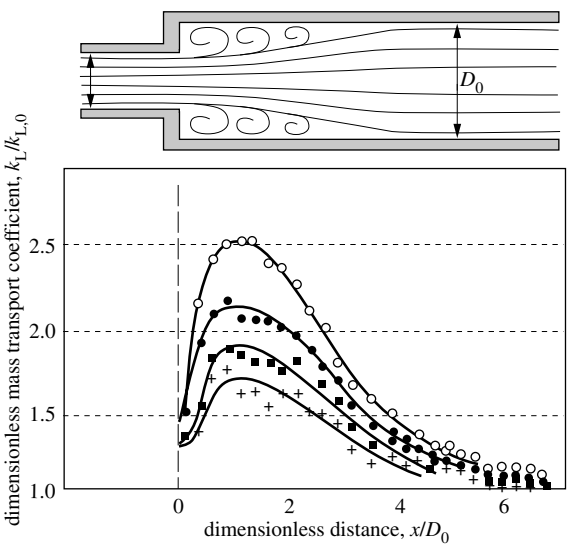


Figure 10.27 Effect of an expansion on local mass transport rate: the dimensionless mass transport coefficient $k_L/k_{L,0}$ is plotted as a function of the dimensionless distance from the expansion x/D_0 . Here $k_{L,0}$ is the mass transport coefficient far away from the expansion and D_0 is the pipe diameter down stream from the expansion. Reynolds number: (o) 2.1×10^4; (•) 4.2×10^4, (■) 8.4×10^4 and (+) 13×10^4. Ratio of pipe diameters: $d_0/D_0 = 0.625$, adapted from [19].

© 2007, first edition, EPFL Press

reduction on a segmented electrode [19]. The highest mass transport rates are observed just downstream from the expansion at a distance of about one pipe diameter. The effect is most pronounced at a Reynolds number of 21000 and becomes smaller at higher Reynolds numbers. Corrosion of copper is highest in the zone of highest mass transport [20].

Influence of surface roughness

The friction coefficient f' for a given flow velocity, increases as the wall rugosity increases and according to (10.53) so does the mass transport rate. As a result, rough surfaces corrode at a higher rate.

Flow accelerated corrosion usually causes an increase in surface roughness. Indeed, copper alloy surfaces corroded under turbulent flow conditions often exhibit characteristic surface features, such as shallow pits elongated in flow direction. As a consequence of increasing surface roughness, the rate of corrosion increases with time. Figure 10.28 schematically shows a possible sequence of events that could cause formation of elongated pits during flow accelerated corrosion of copper in an aerated solution [21]. A flow disturbance causes a locally higher mass transport rate and the higher corrosion rate leads to thinning of the corrosion product layer (a). Eventually bare metal surface is exposed, which further accelerates corrosion (b). The growing recess intensifies the local turbulence, leading with time to the formation of elongated corrosion pits with a characteristic shape (c).

Figure 10.29 shows corrosion current densities measured on the inside wall of a copper pipe in a 3% NaCl solution at 20 °C as a function of the flow velocity. During a ten-minute test (a), the surface remains relatively smooth and the corrosion rate hardly changes with the flow velocity. During a 100-hour test (b), the surface roughness increases significantly and so does the corrosion rate, which, in addition, becomes increasingly dependent on flow velocity. The results show that the critical

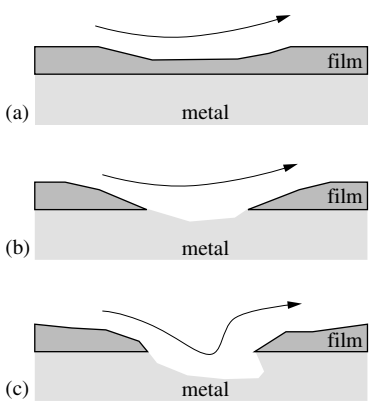


Figure 10.28 Sequence of events in erosion corrosion (schematic): (a) erosion of the film; (b) corrosion of the metallic surface exposed to the flow; and (c) formation of pits with characteristic elongated shape.

© 2007, first edition, EPFL Press

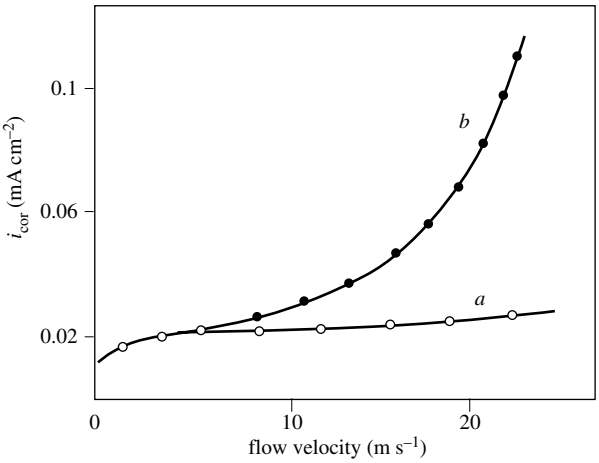


Figure 10.29 Effect of flow velocity on corrosion rate of copper in 3% NaCl: (a) test of 10-minute duration; and (b) test of 100-hour duration [21].

flow velocity leading to accelerated corrosion depends strongly on the surface state of the metal, which evolves with time. The values of Table 10.26 therefore are not precise, they just indicate typical limits observed during prolonged testing under specific conditions.

10.4.2 Erosion corrosion

Erosion due to impingement of particles entrained by a gas

Mechanical damage, rather than chemical effects, normally dominate during erosion caused by impingement of particles carried by a gas under ambient conditions. To assess the effect of different variables on the rate of erosion we consider a simple model.

A hard particle strikes the surface of a ductile metal at a 90° angle and makes an indentation by plastic deformation of the metal (Figure 10.30). Neglecting dynamic effects, we can view this process in the same way as making an indentation in a hardness test. If the hardness is assumed to be constant (no work hardening), the work performed for making the indentation, $W_{F,j}$ is given by:

$$W_{F,j} = \int_0^h F(x)\mathrm{d}x = \int_0^h HA_j(x)\,\mathrm{d}x = HV_j \tag{10.54}$$

Here $F(x)$ is the normal force exerted by the particle, $A_j(x)$ is the cross section of the indentation at a depth x, H is the hardness of the metal and V_j is the volume of the indentation.

© 2007, first edition, EPFL Press

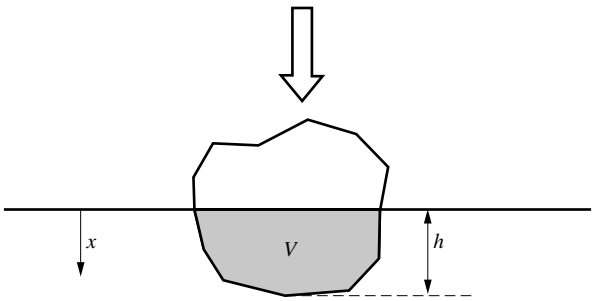


Figure 10.30 Plastic deformation at the surface of a ductile metal resulting from the impact of a hard particle.

The kinetic energy of the impinging particle $E_{\mathrm{kin,j}}$ is

$$E_{\mathrm{kin,j}} = \frac{m_{\mathrm{j}} v_{\mathrm{j}}^2}{2} \tag{10.55}$$

where m_{j} is the mass of the particle and v_{j} its velocity perpendicular to the surface. Only a part of the displaced material is lost as wear particles. Taking this into account by introducing a factor $K_{\mathrm{e,j}}$, we obtain, with equations (10.54) and (10.55), an expression for the volume of metal lost due to the impact of a single particle:

$$V_{\mathrm{e,j}} = K_{\mathrm{e,j}} \frac{m_{\mathrm{j}} v_{\mathrm{j}}^2}{2H} \tag{10.56}$$

The particle flux, N_{p} (particles per second) multiplied by the average mass of the particles, $\mathrm{m_p}$ and the time t gives the total mass of particles impinging during a given time. Defining the ***erosive wear coefficient*** K_{e} as the fraction of deformed material lost by erosive wear in the average, and the average particle velocity, v_{p}, we obtain for the wear volume due to particle impingement

$$V_{\mathrm{e}} = K_{\mathrm{e}} N_{\mathrm{p}} m_{\mathrm{p}} \frac{v_{\mathrm{p}}^2}{2H} t \tag{10.57}$$

Differentiation with respect to time yields the rate of erosive wear v_{e}:

$$v_{\mathrm{e}} = \frac{\mathrm{d}V_{\mathrm{e}}}{\mathrm{d}t} = K_{\mathrm{e}}\, N_{\mathrm{p}}\, m_{\mathrm{p}} \frac{v_{\mathrm{p}}^2}{2H} \tag{10.58}$$

Although the described model is much simplified, it brings out a few interesting aspects of erosive wear of ductile materials. In a similar way as in sliding wear, according to equation (10.58) the rate of erosion should be inversely proportional to hardness. The predicted effect of hardness has been verified for pure metals, but does not well hold for steels for which the hardness depends on the microstructure [22]. The model further predicts that the rate of erosive wear should be proportional to the

© 2007, first edition, EPFL Press

square of the impact velocity of the particles. Experiments confirm this effect, but the exponent is generally found to be slightly higher than two [22].

The presented simple erosion model applies to ductile materials only; erosion of brittle materials involves cracking instead of plastic deformation and requires a different approach. Furthermore, the model does not explain the mechanism of formation and detachment of particles, which involves, in a complex fashion, plowing, cutting and chipping phenomena and depends, among other, on particle shape, materials properties and angle of incidence.

Erosion corrosion of passive metals

In the absence of suspended particles, the corrosion rate of passive metals such as stainless steel or titanium in neutral media is not markedly affected by hydrodynamic conditions (Table 10.26). However, when exposed to slurries, these metals are subject to erosion corrosion because the suspended particles that impinge on the surface damage the passive film. As a consequence an anodic partial current flows which serves for film repair and repassivation of damaged areas. In the presence of aggressive anions such as chloride, passive film damage can lead to metal dissolution by pitting [23].

To study these phenomena it is necessary to carry out experiments that allow one to control the velocity and attack angle of the particles as well as their size, shape and concentration. This can be achieved by the use of an impinging jet of solution with a defined particle concentration. Figure 10.31 shows results of such experiments. A jet of 0.6 M NaCl solution containing sand particles impinges on a 304 stainless steel surface that is kept at a constant potential in the passive region. Each impact produces an acoustic signal and an anodic current transient. The charge underneath the anodic current transients is a measure of the amount of metal oxidized during repassivation

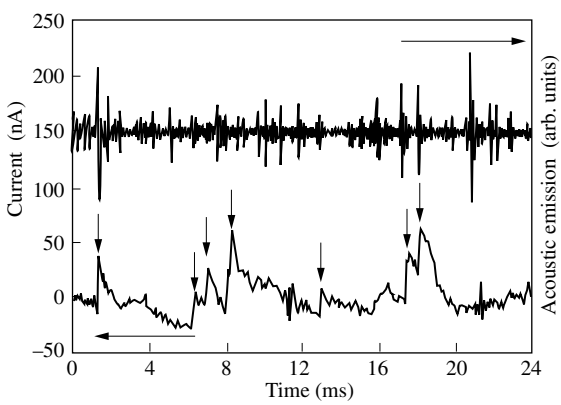


Figure 10.31 Current transients and acoustic emission signals resulting from particle impingement on a AISI 304 stainless steel. A 0.6 M NaCl solution containing sand particles with a diameter of some 200 μm is projected on the passive steel surface held at a potential of –0,06 V vs. SCE, adapted from [23].

© 2007, first edition, EPFL Press

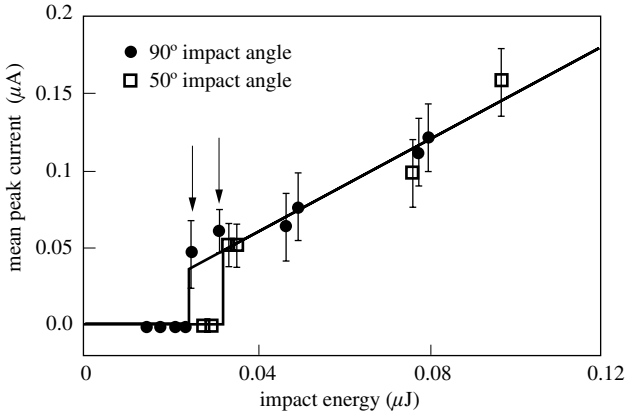


Figure 10.32 Mean transient current resulting from particle impact plotted as a function of the kinetic energy of the particles. A 0.6 M NaCl solution containing sand particles of about 200 μm diameter is projected on a AISI Type 304 stainless steel surface held at a potential of –0.06 V_{SCE} in the passive potential region. Impact angles are 90° and 50° respectively [23].

and hence of the rate of erosion corrosion. While the current transients correlate well with the acoustic emission bursts, there are more acoustic emission bursts than current transients. This is simply because in these experiments particles also impinged on the electrode holder.

The passive film damage caused by an impacting particle of given size and shape should depend on its kinetic energy. The data shown in Figure 10.32 confirm this expectation. They indicate that there is a threshold energy below which the damage of the passive film is negligible. Not surprisingly, the threshold value for given conditions depends on the angle of incidence.

Erosion corrosion in turbulent pipe flow

We have seen in the previous section that under turbulent flow conditions an abrupt enlargement of the pipe diameter leads to a higher mass transport rate just down stream from an expansion of the pipe diameter. This is due to local eddies forming in this region. Solid particles strike the surface with higher energy in this zone, abrading the surface of non-passive metals or damaging the protecting oxide film on passive metals.

As a result the area just downstream from a widening of the pipe diameter will corrode preferentially. Figure 10.33 illustrates this phenomenon. It shows the rate of erosion corrosion of a 15% chromium steel sample in contact with water that contains sand. The rate of erosion corrosion is particularly high in the narrow part of the tube, as well as at a short distance downstream from the expansion. The material loss exhibits a maximum at a relative distance where the fluctuation of the velocity vector perpendicular to the wall is greatest [20]. Under the conditions of the experiments the

© 2007, first edition, EPFL Press

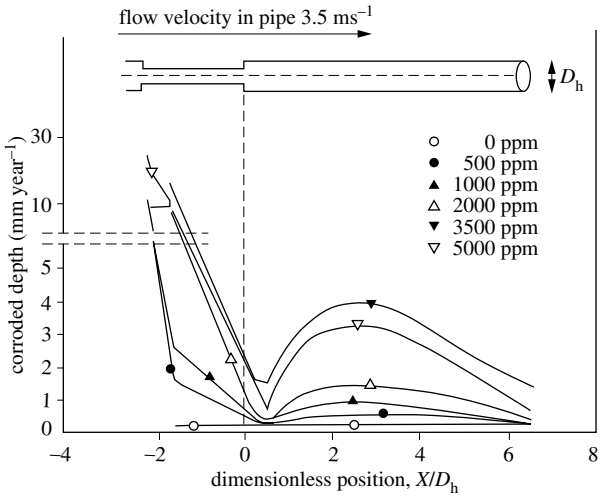


Figure 10.33 Effect of sand content of water on erosion corrosion of Fe-13Cr steel as a function of distance downstream of a widening of a pipe. Flow is from left to right; the flow rate is 3.5 m s^{-1} [19].

maximum rate of erosion corrosion is observed at a distance of about two pipe diameters downstream from the expansion.

The data of Figure 10.33 indicate that the loss of material, at a given position, increases proportionally with the sand content of the water and with the kinetic energy of the particles, which varies as the square of the flow velocity.

$$v_{\text{cor}} \propto c_{\text{p}} E_{\text{kin,p}} \tag{10.59}$$

Here, c_{p} represents the particle concentration and $E_{\text{kin,p}} = (1/2)m_{\text{p}}v_{\text{p}}^2$, the kinetic energy of the particles.

In reality, the rate of erosion corrosion will also depend on the hardness and shape of the particles, the effective impact angle and the mechanical properties of the surface film and the underlying metal.

10.4.3 Cavitation corrosion

Cavitation corrosion refers to the progressive deterioration of a surface under the combined action of corrosion and the implosion of cavitation bubbles. These bubbles can form in fluids undergoing turbulent flow whenever a local low-pressure zone is created; this can occur at a position downstream from an orifice in a pipe, on certain parts of a boat propeller, or in hydraulic turbines.

Origins of degradation by cavitation

In the turbulent flow regime, the local pressure of a fluid fluctuates. If, at certain positions, the pressure drops below the vapor pressure of the fluid, a small amount of

© 2007, first edition, EPFL Press

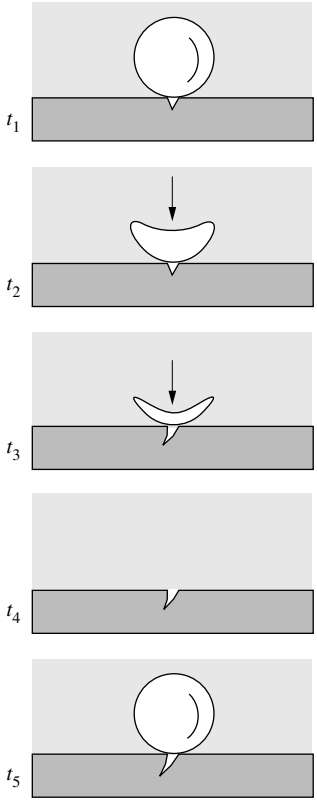


Figure 10.34 Formation and implosion of cavitation bubbles (schematic).

liquid can evaporate, and we observe the formation of a vapor bubble, called a cavitation bubble (Figure 10.34). With diameters that can vary between several micrometer and a millimeter, these bubbles grow from particular sites, related to surface topography or the presence of inclusions.

Because of the pressure fluctuations, the cavitation bubbles implode when the local pressure increases again. This phenomenon produces a shock wave. During a short period of several microseconds, the pressure at the metal surface can thus reach very high values, up to 1000 MPa. The repetition of this mechanism induces fatigue at the metal surface, and thus causes irreversible damage. Attack patterns of characteristic shape appear at the deteriorated zones.

The implosion phenomenon is presented in more detail in Figure 10.35, which gives the result of a theoretical calculation on the evolution of the shape of a cavitation bubble as it implodes: it recedes at its center, and then a liquid jet forms which perpendicularly strikes the metal surface. High-speed photographs of imploding cavitation bubbles confirmed this behavior.

© 2007, first edition, EPFL Press

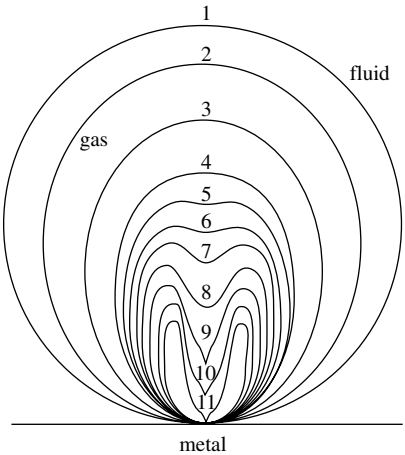


Figure 10.35 Theoretical evolution of the form of an imploding cavitation bubble [24].

Experimental study of cavitation corrosion

Figure 10.36 shows three experimental configurations frequently used to study cavitation corrosion:

- rotating disk in the turbulent-flow regime;
- vibrating device;
- Venturi system.

The ***turbulent-flow rotating disk*** allows the operator to simultaneously study a number of samples. The rotation rate of the disk is easily varied and can attain very high linear velocities at the sample surface. On the other hand, the set up is not well

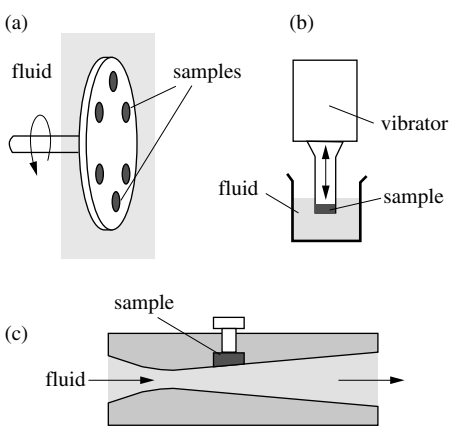


Figure 10.36 Experimental study of cavitation corrosion: (a) rotating disk in the turbulent-flow regime; (b) vibrating device; and (c) Venturi system.

© 2007, first edition, EPFL Press

suited for electrochemical studies. The placement of obstacles on the disk upstream from the samples facilitates the creation of cavitation bubbles.

The ***vibrating device*** functions on the principle of magnetostriction or piezoelectricity. A transducer supplies vibrations at a frequency on the order of 1 kHz, with an amplitude of 20 to 100 μm typically. During each oscillation, cavitation bubbles appear as the sample is pulled upwards, and these then implode as the sample is pushed back downwards. This experimental arrangement stands out by its simplicity; it does not require any heavy equipment and it requires only small volumes of solution. In addition, it is well suited for electrochemical experiments. On the other hand, because the entire surface is exposed to intense cavitation, the application of the results to other flow systems may pose problems.

Venturi systems exist in a variety of forms, but their principle of operation is always the same: a fluid in the turbulent flow regime flows in a pipe and passes through a contraction that generates cavitation bubbles. These are entrained by the fluid and implode downstream at the point where the sample is located. Implosion of cavitation bubbles at this point is induced by an enlargement of the pipe diameter, which causes a deceleration of the fluid velocity and therefore an increase in the local pressure. Among the three systems shown in Figure 10.36, the Venturi system provides experimental conditions that are closest to those found in practice. In addition, it can easily be adapted to carry out electrochemical measurements or to optical observation of the cavitation bubbles. However, Venturi systems require a powerful pumping system and a large volume of solution, which makes them rather cumbersome. Furthermore, to obtain representative results, experiments must often be run for prolonged periods of time.

Interpretation of the results

Because of the large number of parameters that come into play in cavitation corrosion, the interpretation of the results obtained in the laboratory and their application to the design of technical installations poses many problems. The ***cavitation number*** N_K is often used to compare cavitation conditions among experimental configurations of comparable geometry. The cavitation number expresses the ratio between the pressure difference $\Delta P = P - P_{vap}$, and the kinetic energy $E_{kin} = (1/2)\rho v^2$ of the fluid.

$$N_K = \frac{P - P_{vap}}{(1/2)\rho u^2} \tag{10.60}$$

Here, P represents the pressure of the fluid and P_{vap} is its vapor pressure at the temperature of the experiment; u designates the flow velocity and ρ is the density of the fluid. The degree of deterioration due to cavitation is proportional to the difference:

$$\Delta N_K = N_{K,crit} - N_K \tag{10.61}$$

The ***critical cavitation number*** $N_{K,crit}$, which appears in expression (10.61), is an empirical quantity, whose value varies as a function of the geometry of the

© 2007, first edition, EPFL Press

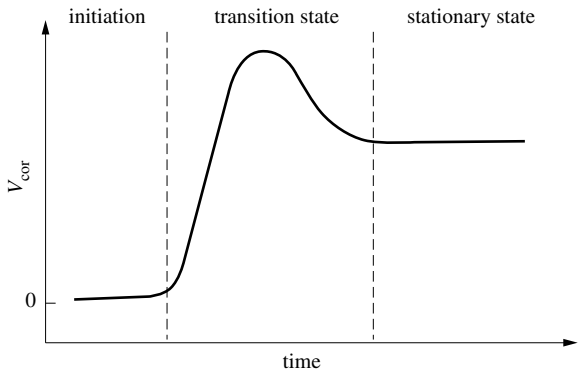


Figure 10.37 Evolution of rate of cavitation corrosion as a function of the test duration (schematic).

experimental apparatus. Cavitation attack can only take place if N_K is smaller than the critical cavitation number, in other words if $\Delta N_K > 1$.

The formation of cavitation bubbles depends on the rugosity, which often changes during the test. The rate of cavitational erosion therefore does not remain constant. One generally observes an onset period, followed by a transition period during which the rate of cavitation corrosion passes through a maximum before it reaches a steady value (Figure 10.37). The duration of these periods depends on the material. To compare the results obtained for different materials the transition period must be short relative to the time scale of the experiment.

Mechanisms of cavitation corrosion

Metal degradation by cavitation is above all a mechanical phenomenon, but corrosion also contributes. Figure 10.38 illustrates this statement. It shows the results of cavitation corrosion tests carried out on steel using the vibrating device described above, in different environments and at different temperatures. The mass loss as a function of temperature passes through a maximum. This behavior can be explained by postulating that two factors, acting with opposite effect, contribute to the deterioration. On one hand, higher temperatures increase the vapor pressure and thus favor cavitation. On the other hand, the solubility of the corrosive agent, in this case oxygen, decreases with temperature, thereby decreasing the corrosion rate. Further evidence for the influence of corrosion is the increased mass loss observed in solutions of greater conductivity.

Cavitation corrosion is still poorly understood. Two mechanisms have been proposed:

- the removal of material by depassivation-repassivation events;
- embrittlement of the metal by hydrogen.

According to the ***depassivation-repassivation*** mechanism the metal loss is due to an oxidation reaction. The implosion of the cavitation bubbles on a passive surface

© 2007, first edition, EPFL Press

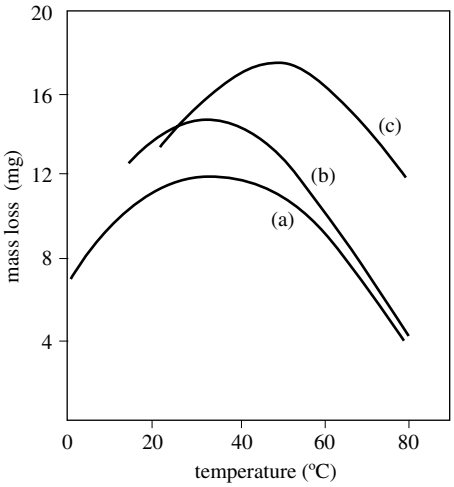


Figure 10.38 Effect of electrolyte on cavitation corrosion. Weight loss of a carbon steel after 15 minutes as a function of temperature during cavitation tests using a vibrating device: (a) in distilled water; (b) in water buffered to pH 8; and (c) in a 3% NaCl solution [25].

creates shock waves that damage the passive film, exposing the underlying metal. Between shocks the surface repassivates and each repassivation event transforms a small amount of metal into oxide. By repeating themselves, these reactions can lead to significant material loss. Figure 10.39 shows the anodic polarization curve of a steel specimen measured in presence and absence of cavitation using a vibrating device [26]. In the presence of vibrations, the passive current density is significantly higher.

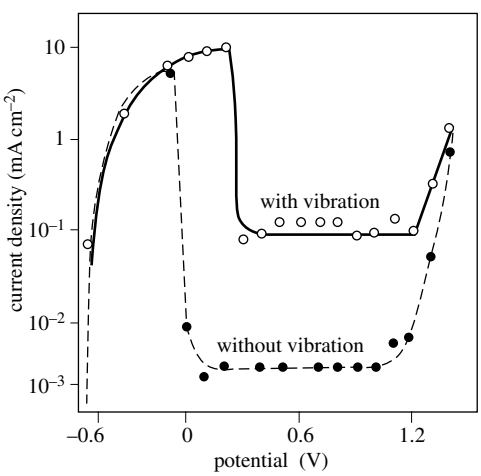


Figure 10.39 Effect of electrode vibration on the anodic polarization curve of steel in a phthalate solution [26].

© 2007, first edition, EPFL Press

This proves that in these experiments cavitation indeed damaged the passive film of the metal.

Cavitation corrosion damage due to ***hydrogen embrittlement*** can be explained as follows. When the surface of an active metal is exposed to an electrolyte, it can react with protons to form hydrogen. Under certain conditions, the hydrogen may diffuse into the metal and increase its brittleness (Chap. 11). This then lowers the fracture energy and facilitates the formation of metal particles due to fatigue cracking under the effect of repeated shock waves from imploding cavitation bubbles.

Somewhat speculative, the proposed mechanisms of cavitation corrosion need further experimental proof. Anodic oxidation, for example, does not explain the specific shape of the attack patterns created by cavitation, which normally exhibit a morphology that is very different from that typically found after anodic dissolution or pitting corrosion. Rather, the morphology of cavitation pits resembles that expected for particle detachment by brittle fracture due to fatigue. Mechanical damage of the metal, rather than of the passive film would then be the cause of cavitation corrosion. A hydrogen embrittlement mechanism could be responsible for such behavior, but it is in contradiction with certain experimental observations. In particular, a cathodic polarization of the metal in the region of hydrogen generation is sometimes applied to suppress cavitation corrosion [25].

To reduce cavitation corrosion damage in cooling circuits, corrosion inhibitors are sometimes added. This underscores the electrochemical nature of cavitation corrosion.

Bibliography

E. Rabinowicz, *Friction and Wear of Materials*, J. Wiley, New York (1965) 244 pp.
I. M. Hutchings, *Tribology*, Edward Arnold, London (1992) 273 pp.
D. H. Buckley, *Surface Effects in Adhesion, Wear and Lubrication*, Elsevier, Amsterdam (1981) 631 pp.
R. B. Waterhouse, *Fretting Corrosion*, Pergamon Press, Oxford (1972) 253 pp.
C. M. Preece, ed., *Erosion, Treatise on Materials Science and Technology,* Vol 16, Academic Press, New York (1979) 450 pp.

References

[1] D. H. Buckley, *Surface Effects in Adhesion, Wear and Lubrication*, Elsevier, Amsterdam (1981) p. 513.
[2] N. P. Suh, *Tribophysics*, Prentice Hall, Englewood Cliffs N. J. (1986) p. 75.
[3] H. S. Kong, M. F. Ashby, *MRS Bulletin* No. 10, 41 (1991).
[4] H. J. Mathieu, D. Landolt , R. Schumacher, *Surf. Interf. Anal*. 9, 477 (1986).
[5] D. Landolt, S. Mischler, M. Stemp, *Electrochimica Acta* 46, 3913 (2001).
[6] J. F. Archard, *J. Appl. Phys.* 24, 981 (1953).
[7] E. Rabinowicz, *Fundamentals of Tribology*, N. P. Suh, N. Saka, eds, The MIT Press, Cambridge Mass. (1978) p. 351.

© 2007, first edition, EPFL Press

[8] E. Rabinowicz, *Friction and Wear of Materials*, J. Wiley, New York (1965) pp. 157, 170, 193.
[9] N. P. Suh, *Tribophysics*, Prentice Hall, Englewood Cliffs, N.J. (1986) p. 282.
[10] S. C. Lim, M. F. Ashby, *Acta Met.* 35, 1 (1987).
[11] T. F. J. Quinn, J. L. Sullivan, D. M. Rowson, *Wear* 94, 175 (1984).
[12] P. Q. Wu, J. P. Celis, *Wear* 256, 489 (2004).
[13] S. Barril, S. Mischler, D. Landolt, *Wear* 256, 963 (2004).
[14] D. Landolt, S. Mischler, M. Stemp, S. Barril, *Wear* 256, 517 (2004).
[15] R. S. Treseder, *NACE Corrosion Engineers Reference Book*, NACE, Houston Texas (1980) p. 68.
[16] K. F. Efird, *Corrosion* NACE 33, 3 (1977).
[17] B. Poulson, *Wear* 233-235, 497 (1999).
[18] M. Nordsveen, S. Nesic, R. Nyborg, A. Stangeland, *Corrosion* 59 (5) 443 (2003).
[19] U. Lotz, M. Schollmaier, E. Heitz, *Werkstoffe und Korrosion* 36, 163 (185).
[20] E. Heitz, *Corrosion* 47, 135 (1991).
[21] C. Loss, E. Heitz, *Werkstoffe und Korrosion* 24, 38 (1972).
[22] I. M. Hutchings, *Tribology*, Edward Arnold, London (1992) 273 pp.
[23] K. Sasaki, G. T. Burstein, *Phil. Mag. Lett.* 80, 489 (2000).
[24] K. A. Morch, C. M. Preece eds., *Erosion, Treatise on Materials Science and Technology* vol. 16, Academic Press, New York (1979) p. 309.
[25] C. M. Preece, *Erosion, Treatise on Materials Science and Technology,* vol. 16, Academic Press, New York (1979) p. 249.
[26] A. Karimi, J. L. Martin, *Int. Mat. Rev.* 31, 1 (1986).

© 2007, first edition, EPFL Press

CHAPTER 11

ENVIRONMENT INDUCED CRACKING

11.1 INTRODUCTION

In certain environments, metals crack or rupture at stress levels significantly lower than their ultimate tensile strength, or even their elastic limit. The phenomenon is known as ***environment induced cracking*** (*EIC*). The terms *environment sensitive cracking* or *environment assisted cracking* are also sometimes used in the literature. To some extent, *EIC* phenomena are also observed with ceramics and polymers, but they are in general less of a problem and the underlying mechanisms differ from those found with metals. The particular importance of *EIC* of metals is due to the fact that load bearing structures for critical applications are usually made of metal. Indeed, *EIC* of metals is a major concern in many technological areas, including chemical and petroleum industries, nuclear power generation, aero-space and civil engineering. In all these applications, an unexpected rupture can lead to accidents with catastrophic consequences for human life and the environment. Design engineers must therefore take into account the fact that if a structure is exposed to corrosion conditions that favor crack formation and growth its mechanical resistance will significantly diminish with time.

Environment induced cracking of metals includes the following phenomena:

- stress corrosion cracking;
- hydrogen embrittlement;
- corrosion fatigue.

Stress corrosion cracking (*SCC*) is caused by the combined action of a static tensile stress and corrosion. The stress may originate from external loading or be the result of a thermal or mechanical treatment that generates residual stresses in the material. Stress corrosion cracking failures usually occur with some delay and they are limited to particular alloy–environment combinations, temperature ranges and stress levels.

Hydrogen embrittlement designates a premature failure of a metal or alloy caused by the absorption of hydrogen. In certain cases, stress corrosion cracking is accompanied by hydrogen embrittlement at the crack tip and it is then difficult to separate the two phenomena.

Corrosion fatigue is the result of the combined action of cyclic loading and corrosion. It manifests itself as a decrease in the endurance limit of the metal.

© 2007, first edition, EPFL Press

Liquid metal embrittlement is also a cause of cracking and premature failure. One could look at it as a special type of *EIC* that may occur when a metallic structure is in contact with liquid metal or with a solid metal that is close to its melting point. Damage due to liquid metal embrittlement has been observed during brazing or in cooling circuits containing a liquid metal. Table 1.1 summarizes the susceptibility of different alloys to this type of damage. Liquid metal embrittlement is essentially governed by surface tension effects at grain boundaries and does not involve chemical reactions with the environment. It shall therefore not be considered further here.

Table 11.1 Sensitivity of some alloys (x = sensitive, o = not sensitive) to liquid metal embrittlement [1]. The melting point of the metal causing embrittlement is given in parentheses.

	Hg (–38 °C)	Sn (232 °C)	Bi (271 °C)	Pb (327 °C)	Zn (419 °C)
Al alloys	x	x	x	x	x
Steels	o	x	o	x	x
Cu alloys	x	x	x	x	o
Ti alloys	o	o	o	o	o

Failure analysis

Fractures resulting from environment induced cracking share two common characteristics:

- they are of the brittle type;
- they occur only after a certain period of time.

The occurrence of a brittle fracture (absence of plastic deformation) in a normally ductile metal is a strong indication for the presence of *EIC*. In failure analysis this is used as a diagnostic criterion. Because *EIC* involves corrosion, failure does not take place immediately, but after a certain time only during which a crack forms and grows to a critical size. This behavior complicates the study of *EIC* in the laboratory where the phenomena must be accelerated and the results then extrapolated to more realistic conditions.

Intergranular and transgranular cracking

Depending on conditions, *EIC* leads to the growth of intergranular cracks, which follow the grain boundaries, or of transgranular cracks, which pass through the grains (Figure 11.2). The optical micrographs shown in Figure 11.3 illustrate the two types of cracks. Photograph (a) shows the intergranular cracks formed in a stainless steel exposed to an alkaline environment at high temperature and photograph (b) shows transgranular cracks with branches that developed in an austenitic stainless steel during exposure to a chloride-containing environment.

The growth of cracks eventually leads to ***intergranular fractures*** or ***transgranular fractures***, depending on the nature of the crack. Sometimes, mixed fractures are observed that partly follow the grain boundaries, partly traverse the grains. Figure

© 2007, first edition, EPFL Press

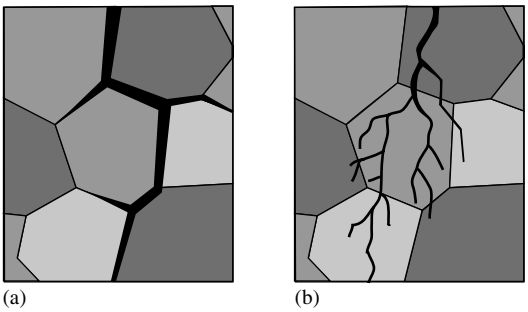


Figure 11.2 Stress corrosion cracking (schematic): (a) intergranular cracking ; (b) transgranular cracking.

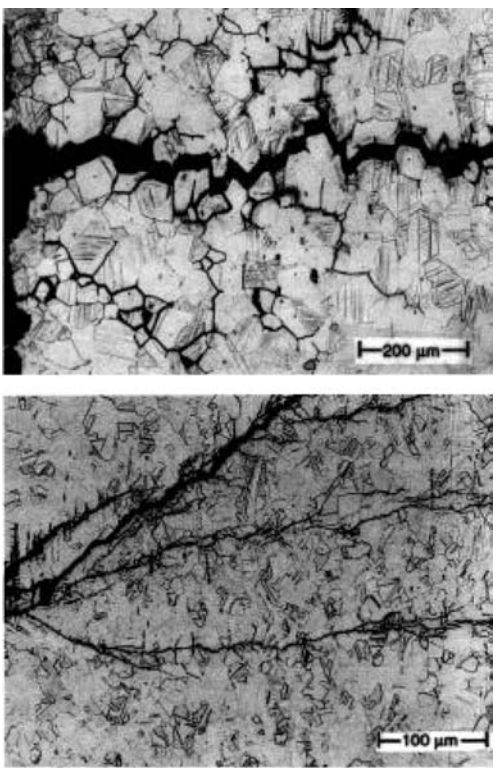


Figure 11.3 Stress corrosion cracks: (a) intergranular cracks of an austenitic stainless steel (DIN 1.4571) exposed to an alkaline environment at 80 °C; (b) transgranular cracks of an an austenitic stainless steel (DIN 1.4541) exposed to a chloride-containing environment at 100 °C [2].

© 2007, first edition, EPFL Press

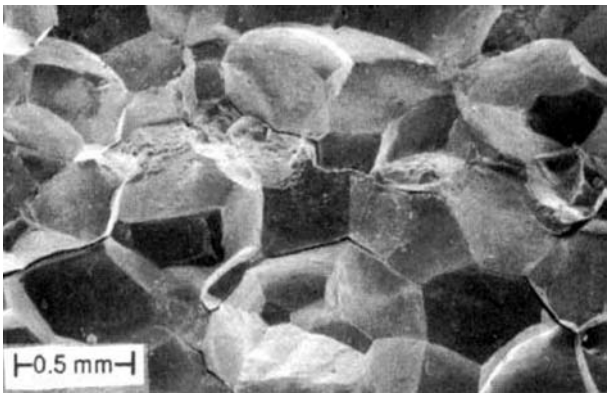


Figure 11.4 Intergranular fracture of Al-4% Cu alloy subjected to stress corrosion cracking test in NaCl.

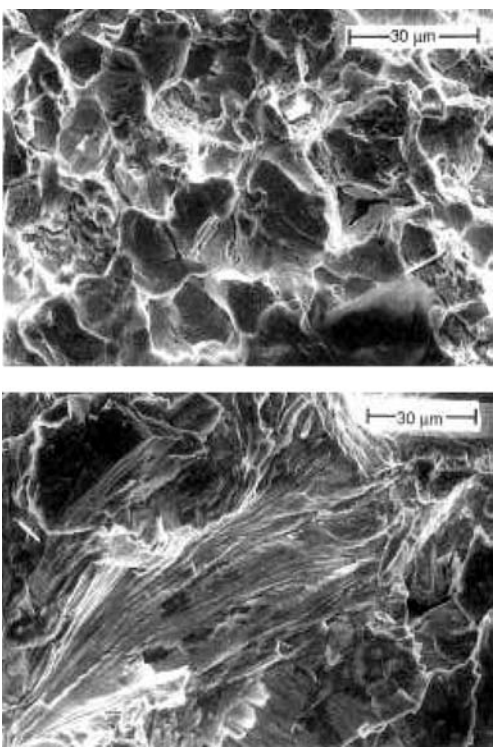


Figure 11.5 Fracture surfaces from stress corrosion cracking: (a) intergranular fracture of carbon steel in an alcaline environment at 70 °C; (b) transgranular fracture of stainless steel (DIN 1.4439) in chloride medium at 100 °C.

© 2007, first edition, EPFL Press

11.4 shows the intergranular fracture due to the stress corrosion cracking of an Al-Cu alloy that had been sensitized by a thermal treatment. The fracture shown in Figure 11.5(a) is also intergranular, but the individual grains are less visible than in Figure 11.4. It represents a stress corrosion cracking failure of a carbon steel in an alkaline medium at elevated temperature. Figure 11.5(b) shows the fracture surface of an austenitic stainless steel exposed to a chloride-containing environment. It exhibits an unusual morphology that resembles that of leaves and the fracture is predominantly transgranular. In practice, the study of fracture surfaces provides precious information on the origin and mechanism of material failures.

11.2 EXPERIMENTAL STUDY OF ENVIRONMENT INDUCED CRACKING

11.2.1 Overview of test methods

The experimental study of *EIC* mechanisms and the development of test methods for life prediction poses a number of problems. First of all, both the electrochemical and the mechanical conditions at the crack tip must be controlled. Control of electrochemical conditions at the crack tip is particularly difficult because of the presence of concentration gradients and of ohmic potential drops in the electrolyte filling the crack. As a consequence of these, the concentration of ions, the pH and the electrode potential at the crack tip may significantly differ from those outside and one must rely on theoretical models to estimate their value. A further complication arises from the fact that, in order to limit the time of experiments, crack growth must be accelerated, for example by employing a higher temperature or a particularly aggressive solution. Results of such experiments are only significant if the more corrosive conditions do not change the mechanism of crack nucleation and growth. From a mechanical point of view a complication results from the fact that the local stress at the crack tip may change during a test because it depends on the crack geometry. The problem is most pronounced if branching of cracks occurs. All these mentioned factors complicate the mechanistic interpretation of laboratory experiments and the extrapolation of accelerated test results to realistic operating conditions. To improve the reliability of predictions different test methods are sometimes used concurrently.

To study environment induced cracking one controls either the load or the strain. In stress corrosion cracking, we thus distinguish ***constant load tests*** and ***constant strain tests***. As an alternative, one may impose a load or a strain that varies during a test. ***Constant strain rate tests*** are often used for the study of stress corrosion cracking. In this method the strain is increased at a slow constant rate until fracture occurs. For the study of corrosion fatigue, the applied stress or strain vary in a cyclic fashion (Section 11.5). To control the mechanical parameters in EIC experiments one normally uses tensile testing machines, but for the application of a constant load or a constant strain in routine experiments much simpler arrangements are possible as will be shown in the next section.

© 2007, first edition, EPFL Press

An important aspect of *EIC* experiments concerns the shape of the test specimens and whether they contain a crack or not at the start of an experiment. One distinguishes two broad categories: ***precracked specimens*** and ***non-precracked specimens***. Precracking is accomplished by subjecting the test specimen to a cyclic stress by means of a tensile test machine in order to initiate a crack. The precracked specimen is then subjected to the chosen test conditions for the study of the rate of crack growth. In non-precracked specimens both crack initiation and crack growth occur during the actual test.

11.2.2 Static tests with non-precracked specimens

Static tests carried out with non-precracked specimens allow us to characterize, for a given environment, the sensitivity of an alloy to stress corrosion cracking and hydrogen embrittlement. Typically, the test specimen subjected to a constant load or a constant strain is exposed to a corrosive environment and one measures the time to failure. The test specimens sometimes contain a notch—not to be confused with precracking—that fixes the location where failure will occur. In constant strain tests, the ***time to failure*** corresponds to the appearance of the first cracks. This kind of experiment thus indicates the crack initiation time. In constant load tests, on the other hand, the time to failure is the time leading to fracture of the specimen. It corresponds to the sum of the crack initiation and the crack propagation times.

Figure 11.6 schematically shows the evolution of the nominal stress (σ_{nom} = applied load/initial cross section) and of the real stress (σ_{re} = applied load/real cross section) as a function of immersion time for *SCC* tests performed at constant strain and at constant load, respectively. During a constant strain experiment, the real stress decreases as the crack propagates. Once the stress becomes sufficiently small, the crack stops before rupture occurs. In a constant load test, on the other hand, the real

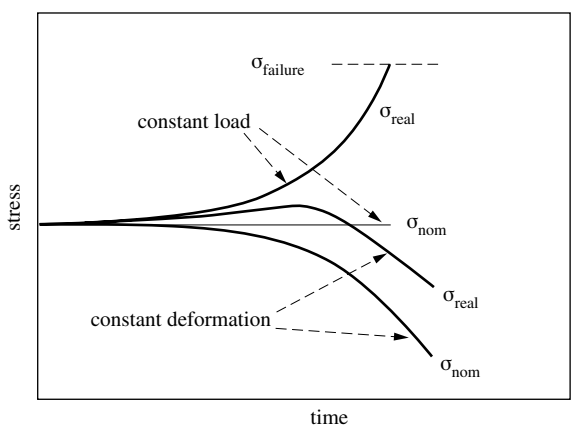


Figure 11.6 Real and nominal stress versus time in stress corrosion cracking tests at constant load and constant deformation.

© 2007, first edition, EPFL Press

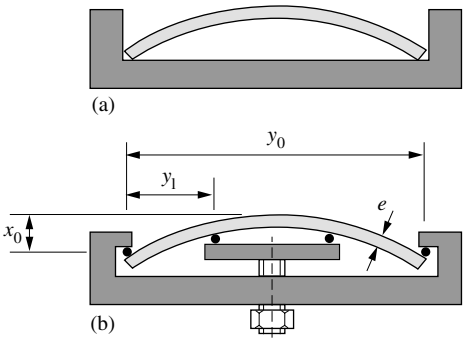


Figure 11.7 Bending tests for stress corrosion cracking.

stress increases as the crack progresses because the cross section decreases, and as a consequence the sample breaks.

A great variety of experimental configurations are used in practice to perform *SCC* tests on non-precracked samples [3]. Figure 11.7 shows two arrangements to carry out ***bending tests*** at constant strain on metal sheets that are not too thick. When the load is applied at two points (Figure 11.7(a)), the tensile stress is maximum at the center of the convex surface and tends towards zero at the two extremities. With the load applied at 4 points (Figure 11.7(b)), the convex surface is subjected to constant stress, σ_s, between the inside points of contact. For small elastic deformations ($x/y_0 \ll 1$), this stress is calculated with the following formula [3]:

$$\sigma_S = \frac{12\,E\,x_0\,e}{3\left(y_0^2 - 4\,y_1^2\right)} \tag{11.1}$$

where E represents the Young's modulus, e is the thickness of the sample, and x_0 is the distance of maximum deformation (Fig. 11.6). As for y_0 and y_1, these designate the distances between the exterior and interior contact points, respectively.

Constant strain bending tests have the advantage of being very simple. They do not require the use of heavy mechanical equipment. By varying the deformation, the tensile stress applied to the convex surface is modified, which allows one, in principle, to study the influence of the stress on the time to failure. However, the stress is not uniform over the entire thickness of the sample: the zone adjacent to the convex surface undergoes tension, and that adjacent to the concave surface is subjected to compression. Bending tests therefore are limited to qualitative studies of *SCC* susceptibility. Because the crack initiation time is not well reproducible, a statistical interpretation of the results may be necessary.

Figure 11.8 shows other test configurations that are often employed. Stressed ***U-bend specimens*** (Figure 11.8(a)) are subjected to a constant plastic deformation. Their use is therefore limited to ductile materials. Because of extensive plastic deformation, fracture occurs in a short time. Qualitative results can thus be obtained rapidly. ***C-ring specimens*** shown in Figures 11.8(b) and (c) are machined from metal

© 2007, first edition, EPFL Press

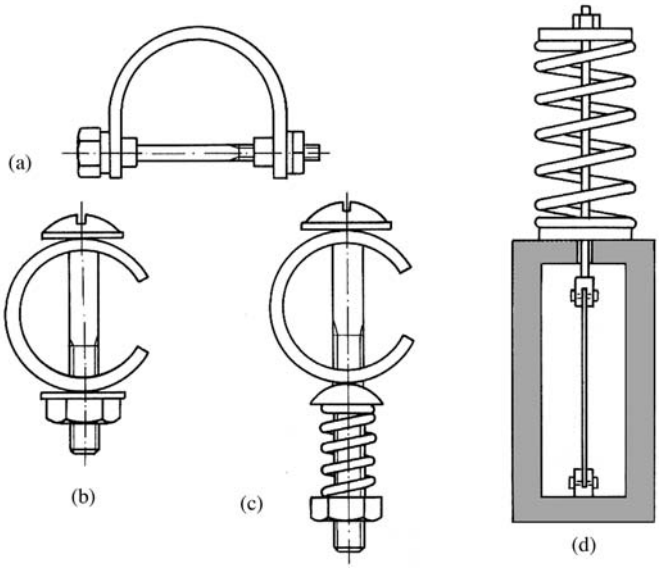


Figure 11.8 Specimens used for the testing of stress corrosion cracking: (a) U-bend specimens; (b) and (c) C-ring specimens; (d) uniaxal-tension-test specimen.

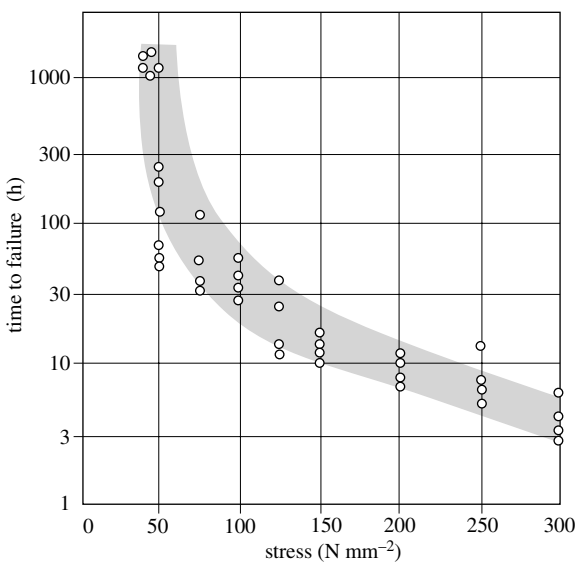


Figure 11.9 Time to failure versus applied stress measured in constant load tests for a type 18-10 austenitic stainless steel specimen. The stress was applied in a boiling solution of 42% $MgCl_2$ [4].

© 2007, first edition, EPFL Press

sheets, bars or tubes. They are used to measure both the crack initiation time and the fracture time.

Uniaxial tension tests offer better control of the load conditions. The specimens that usually have a circular cross section (bars, wires), are subjected to a constant load by use of a spring (Figure 11.8(d)). The time to failure is then measured at different applied loads. A tensile testing machine gives the same information, but is more costly. Figure 11.9 shows the time to failure as a function of applied nominal stress for an austenitic stainless steel, immersed into a boiling solution of 42% $MgCl_2$. Below a certain threshold, called the ***threshold stress for stress corrosion cracking***, no cracking takes place.

11.2.3 Static tests with precracked specimens

Principle of the method

The use of precracked test specimens allows one to measure the crack propagation rate and to determine the threshold stress below which a crack is not able to propagate. On the other hand, this type of test provides no information on the crack initiation time or on the influence of the surface state on stress corrosion cracking.

For most tests carried out with non-precracked specimens, the results depend on the shape and size of the test specimens used. In contrast, the results obtained with precracked samples can be interpreted in terms of ***linear elastic fracture mechanics***, rendering them directly applicable by construction engineers. However, as this method neglects the time for crack initiation, it tends to underestimate the lifetime of a given structure.

Linear elastic fracture mechanics provides a relation between the applied force and the local stress at the crack tip in a brittle material. For a sample under *mode-I* loading (Figure 11.10), the three components of the local stress in the vicinity of the crack tip vary as a function of the radial distance r and the angle θ, the latter taken with respect to the plane of cracking (Figure 11.11). The y component of stress σ_y is given by:

$$\sigma_y = K_I f_{(r,\theta)} \tag{11.2}$$

Here K_I is the ***stress intensity factor***. The subscript *I* indicates the loading mode. For a plate of infinite size, which has a pointed crack of length a on its surface

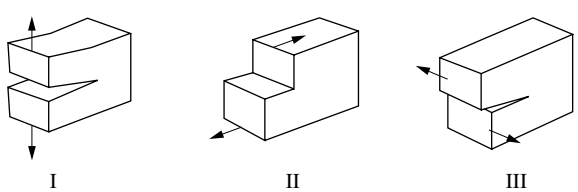


Figure 11.10 Fracture modes I, II, and III.

© 2007, first edition, EPFL Press

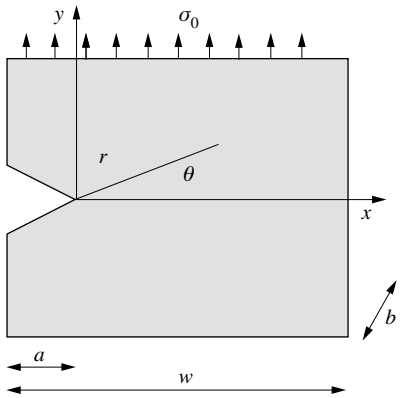


Figure 11.11 Coordinate system with origin at crack tip.

perpendicular to the direction of the applied load (plane strain), the stress intensity factor K_I (N m$^{-3/2}$) is equal to:

$$K_I = \sigma_0 (\pi a)^{1/2} \tag{11.3}$$

where σ_0 (N/m^2) represents the applied stress. For test specimens of finite size, the value of K_I depends on their geometry, in particular on the ratio a/w, where w designates the width. In this case, we must replace (11.3) by (11.4), with the factor $f_{(a/w)}$ expressing the influence of the specimen geometry.

$$K_I = \sigma_0 (\pi a)^{1/2} f_{(a/w)} \tag{11.4}$$

In the absence of corrosive agents, a crack will not propagate unless the stress intensity factor exceeds a critical value, K_{IC}, called the *plain strain fracture toughness*,

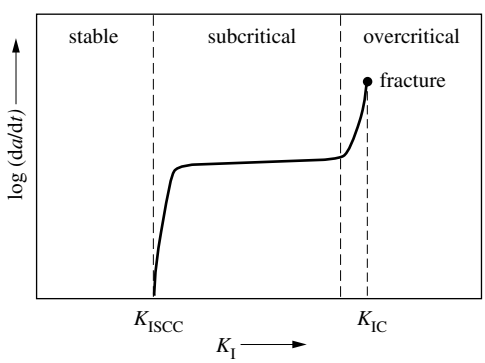


Figure 11.12 Variation of crack propagation rate with stress intensity factor during a stress corrosion cracking test (schematic).

© 2007, first edition, EPFL Press

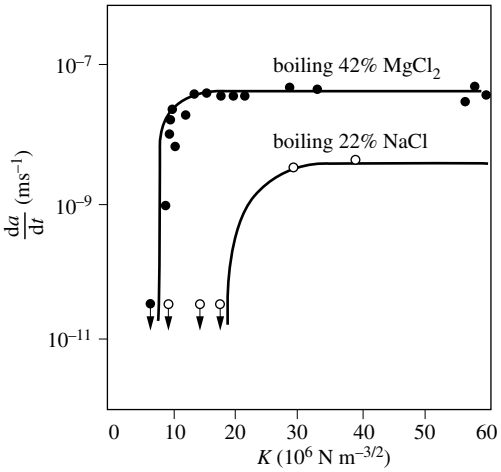


Figure 11.13 Crack-propagation rate versus stress intensity factor for a type-304 austenitic stainless steel: (o) boiling 22% NaCl solution (105 °C); (•) boiling 42% MgCl solution (130 °C).

or, more simply, ***the fracture toughness.*** The fracture toughness is a materials property that characterizes the fracture energy.

In corrosive environments, cracks will sometimes propagate, even though K_I is significantly smaller than K_{IC}. The ***threshold stress intensity factor for stress corrosion cracking***, K_{ISCC}, corresponds to the minimum value of K_I that allows cracks to grow in a given corrosive environment. The experimental value of K_{ISCC} depends, to some degree, on the sensitivity of the method that is used to detect the movement of cracks. The smallest detectable crack growth rates using current methods are on the order of 10^{-10} m/s.

Figure 11.12 shows the crack growth rate (logarithmic scale) as a function of K_I. There are three distinct domains. As long as K_I is smaller than K_{ISCC} , the growth rate is practically zero. In this region (stable region) the alloy is not sensitive to stress corrosion cracking. Above K_{ISCC}, the crack growth rate increases sharply, and then stabilizes at a constant value that depends on the environmental conditions (subcritical region). When K_I reaches K_{IC}, the growth rate accelerates rapidly until fracture occurs (overcritical region). As an illustration, the Figure 11.13 shows the rate of crack growth in an austenitic stainless steel exposed to a 42% $MgCl_2$ solution at 130 °C, or to a 22% NaCl solution at 105 °C. In the more corrosive solution ($MgCl_2$), the crack growth rate is significantly higher, demonstrating the critical role of the corrosive environment.

Geometry of test specimens

The critical dimensions of precracked *EIC* test specimens are defined in various standards [3]. Figure 11.14 presents two types of test specimens that have been widely employed:

© 2007, first edition, EPFL Press

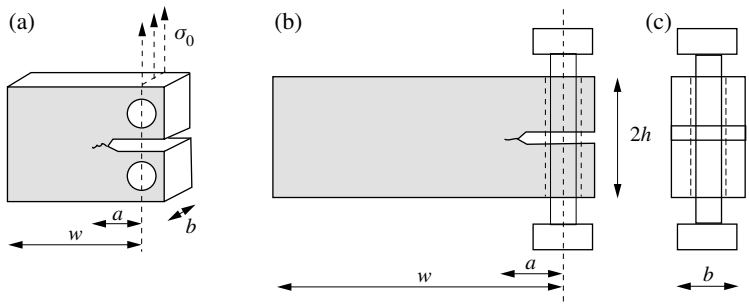


Figure 11.14 Pre-cracked probe shapes for stress corrosion cracking tests: (a) compact tension specimen (CT); and (b) double cantilever beam specimen (DCB).

- compact tension specimens (CT);
- double cantilever beam specimens (DCB).

Compact tension specimens are often used in experiments carried out at constant load. In this case, the stress intensity factor is corresponds to

$$K_{\mathrm{I}} = f_{(\mathrm{a/w})} \frac{P}{bw^{1/2}} \tag{11.5}$$

Here, P designates the applied load ($P = \sigma_0\ wb$), and $f_{(\mathrm{a/w})}$ represents a dimensionless factor, which for $0.2 < a/w < 1.0$ is given by:

$$f_{(\mathrm{a/w})} = \frac{2 + \dfrac{a}{w}}{\left(1 - \dfrac{a}{w}\right)^{3/2}} \left[0.886 + 4.64\,\frac{a}{w} - 13.32\left(\frac{a}{w}\right)^2 + 14.72\left(\frac{a}{w}\right)^3 - 5.6\left(\frac{a}{w}\right)^4 \right] \tag{11.6}$$

Double cantilever beam specimens are widely applied for constant strain testing. If their width w is at least 10 times greater than their height h, the stress intensity depends only on h, and not on w. In addition, when $2 < a/h < 5$, K_{I} is given by the following simple relation [3]:

$$K_{\mathrm{I}} = \frac{E\,\Delta y\,h\left[3h(a+0.6h)^2 + h^3\right]^{1/2}}{4\left[(a+0.6h)^3 + ah^2\right]} \tag{11.7}$$

Here, Δy represents the strain applied in the plane of the screw as indicated in Figure 11.14.

Variation of the stress intensity during crack growth

The stress intensity in front of a crack changes during an experiment as the crack grows. In *constant load* experiments, K_{I} increases with time, because as the crack

© 2007, first edition, EPFL Press

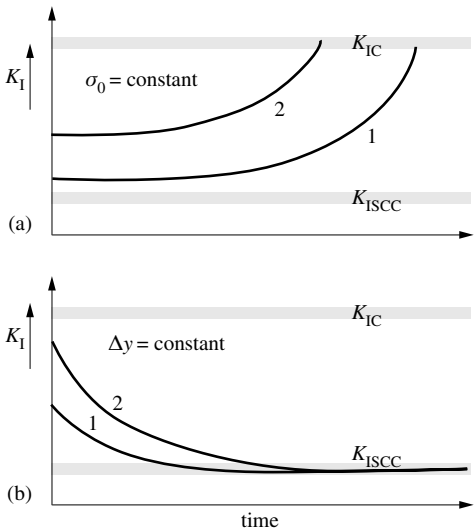


Figure 11.15 Variation of stress intensity at crack tip as a function of the testing time: (a) constant load; and (b) constant deformation. The labels 1 and 2 refer to two tests that differ by the magnitude of the applied load or deformation.

grows the real cross section of the sample decreases (Figure 11.15(a)). In this case, only a single sample is required to measure the curve $da/dt = f(K_I)$ in the subcritical region. On the other hand, several specimens are generally required to determine K_{ISCC}.

In *constant strain* experiments, K_I decreases as a function of time, because the effective stress decreases. The crack stops when the stress intensity reaches the value of K_{ISCC} (Figure 11.15(b)). A single experiment permits one to measure K_{ISCC} and the subcritical crack propagation rate. On the other hand, a series of tests is needed if K_{IC} is to be determined.

Both types of experiments allow us to obtain, with a limited number of samples, quantitative information about the rate of crack growth as a function of stress intensity. On the other hand, the measurement of very slow crack growth rates required to determine K_{ISCC} is quite delicate.

Influence of the plasticity

Linear elastic fracture mechanics is only applicable if the extent of the ***plastic zone*** at the crack tip is negligible compared to the sample size. The presence of a plastic zone at the crack tip can be explained in the following way: the component σ_y of the stress in the plane of the crack (Fig. 11.11) near the tip is equal to:

$$\sigma_y = \frac{K_I}{(2\pi r)^{1/2}} \tag{11.8}$$

© 2007, first edition, EPFL Press

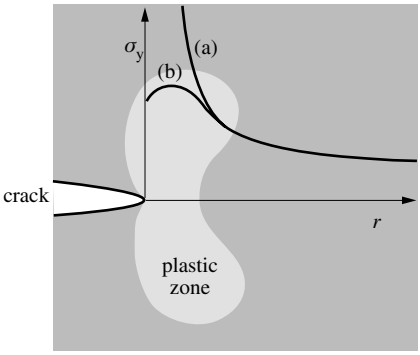


Figure 11.16 Effect of plastic deformation on stress intensity at crack tip: (a) in the absence and (b) in the presence of a plastic deformation zone (schematic).

According to this formula, σ_y should be infinite at the crack tip, in other words, as the radial coordinate r tends to zero (Figure 11.16, curve *a*). In reality, the region of the crack tip undergoes plastic deformation which limits the maximum stress to a value near the yield strength of the material. The exact value depends on the geometry of the sample and on the deformation behavior of the material. In general, the plastic zone is considered to be negligible with respect to the crack size and sample dimension if the following conditions are met:

$$a > 2.5\,(K_I\,/\sigma_e)^2 \tag{11.9}$$
$$b > 2.5\,(K_I\,/\sigma_e)^2 \tag{11.10}$$

Here σ_e represents the yield strength of the metal, and the dimensions a and b are defined in Figure 11.14. The conditions (10.9) and (10.10) indicate the minimum size of precracked specimens. To the extent that these are met, the results of stress corrosion cracking experiments can be interpreted using fracture mechanics.

11.2.4 Slow strain rate testing

The *slow strain rate test* is based on the postulate that the initiation and growth of cracks in a corrosive environment is governed by the strain rate rather than by the stress intensity at the crack tip. The latter plays only an indirect role by causing the local strain.

The test specimen, which may be smooth or notched, is fixed in a tensile testing machine permitting to program the strain rate. A slow constant strain rate is applied and maintained throughout the experiment until the specimen fractures. During the test one measures the stress strain curve, $\sigma = f(\varepsilon)$ and after the test one studies the cross sectional area and the topography of the fracture surface.

© 2007, first edition, EPFL Press

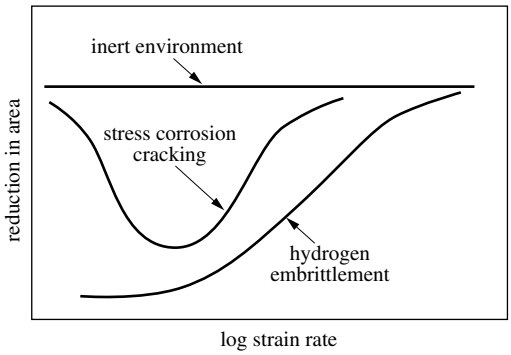


Figure 11.17 Effect of strain rate on the striction coefficient in constant deformation tests in inert surroundings, in the presence of stress corrosion cracking, and in the presence of hydrogen embrittlement.

The ***reduction in area*** k_A expresses the difference between the initial cross sectional area A_0 of the specimen and the fracture area A measured after rupture:

$$k_A = \frac{A_0 - A}{A_0} \tag{11.11}$$

For a brittle fracture, $k_A = 0$ because the fracture area is equal to the initial cross sectional area. For a ductile fracture, on the other hand, $A < A_0$ and therefore $0 < k_A < 1$. To evaluate the influence of corrosion, we compare the values of k_A obtained in an aggressive environment with that observed in an inert environment. Stress corrosion cracking and hydrogen embrittlement favor the occurrence of brittle fractures and therefore they manifest themselves by a smaller value of k_A compared

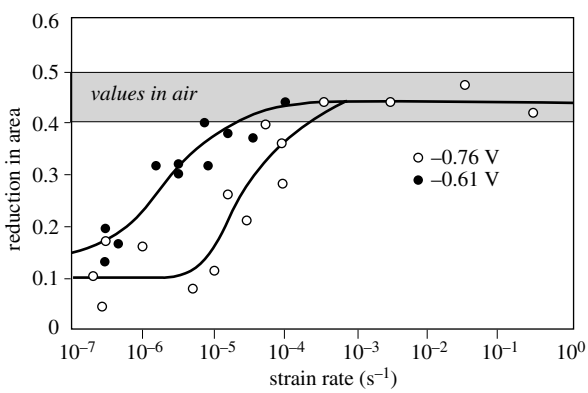


Figure 11.18 Reduction in area as a function of the strain rate for type AISI 4340 steel subjected to cathodic polarization in seawater leading to hydrogen embrittlement: (o) polarization of 0.76 V; (•)polarization of 0.61 V. The grey zone shows values typically found in air.

© 2007, first edition, EPFL Press

to mechanical rupture. The importance of this effect depends on the applied strain rate.

The behavior is schematically represented in Figure 11.17. The influence of corrosion is most important at strain rates of between 10^{-3} and 10^{-7} s^{-1}. In practice, for routine tests, a value of 10^{-6} s^{-1} is often employed. If the strain rate is too high, the corrosion will not have time to act and the metal will behave as in a mechanical test performed in an inert environment. If, on the other hand, the strain rate is too low, corrosion may increase the radius of the crack tip thereby decreasing the stress intensity. As a consequence, the crack may stop growing.

The described behavior applies only if crack propagation involves anodic dissolution at the crack tip. If local hydrogen embrittlement is the cause of cracking, there is no lower limit, since the crack tip is not blunted by corrosion. Figure 11.18 shows experimental measurements taken on a steel specimen, subjected to cathodic polarization at two different potentials, in artificial seawater. The results indicate that hydrogen embrittlement of the metal occurs under these conditions.

The reduction in area observed in slow strain rate experiments in corrosive environments is equivalent to a decrease of the ductility of the metal. Other measurable parameters also depend on the ductility, namely the extent of the plastic zone of the stress strain curve, which determines the *elongation at fracture,* and the area under the nominal stress strain curve, which is a measure for the *fracture energy*. These parameters can also be used for the determination of the sensitivity of an alloy to stress corrosion cracking. Figure 11.19 shows the stress strain curve of cold rolled carbon steel, measured in air and in a NaCl solution containing H_2S, respectively. Hydrogen embrittlement taking place in the NaCl solution induces a significant decrease of both the elongation at fracture and the area under the stress-strain curve.

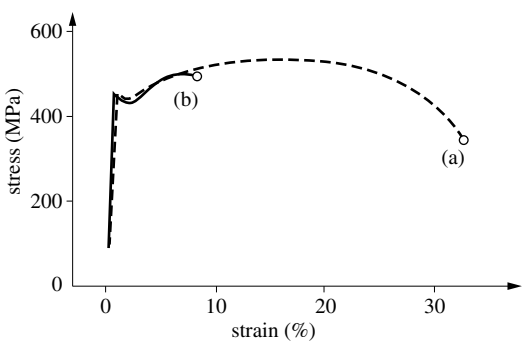


Figure 11.19 Stress-strain curve of cold rolled steel with and without hydrogen embrittlement: (a) stress-strain curve measured in air with a strain rate of 10^{-3} mm s^{-1}; (b) stress-strain curve measured in seawater in the presence of H_2S at a strain rate of 10^{-4} mm s^{-1}.

© 2007, first edition, EPFL Press

11.3 HYDROGEN EMBRITTLEMENT

11.3.1 Solubility of hydrogen in metals

Sources of hydrogen

Hydrogen can dissolve in metals and modify their properties. This phenomenon is a frequent cause of damage, as it does not require that the metal be directly exposed to gaseous hydrogen; most often the hydrogen forms at the metal surface itself by an electrochemical reaction. Three sources of hydrogen can be identified that lead to embrittlement of metals:

- high pressure hydrogen gas;
- hydrogen formed electrochemically by cathodic polarization;
- hydrogen formed during a corrosion reaction.

Certain processes in the chemical and petroleum industries use gaseous hydrogen at high-pressure. Under these conditions, some hydrogen may dissolve into the walls of autoclaves and piping systems made of metal and damage them.

Cathodic hydrogen evolution is a frequent side reaction observed in metal plating. If some of this hydrogen dissolves in the metal forming the cathode it causes embrittlement. This is a well know problem in the fabrication of screws and bolts made of high strength steel and electroplated with zinc. Hydrogen that is formed during electrolytic cleaning or during cathodic protection of steel (Chap. 12) can also lead to embrittlement.

Corrosion reactions taking place in acidic environments give off hydrogen. Similarly, chemical surface treatments such as pickling of steel, or phosphatizing under acidic conditions produce hydrogen at the metal surface.

Thermodynamic solubility

When dissolving into a metal, hydrogen molecules dissociate into atoms:

$$H_2 = 2\,H_m \tag{11.12}$$

In this equilibrium, H_2 represents gaseous molecular hydrogen, and H_m designates the atomic hydrogen dissolved in the metal. The equilibrium constant is written as:

$$K = \frac{X_H}{P_{H2}^{1/2}} \tag{11.13}$$

where X_H corresponds to the mole fraction of hydrogen in the metal and P_{H2} is its partial pressure. The standard free energy of reaction (11.12) is equal to:

$$\Delta G^0 = \Delta H^0 - T\,\Delta S^0 = -RT\,\ln\frac{X_H}{P_{H2}^{1/2}} \tag{11.14}$$

© 2007, first edition, EPFL Press

Table 11.20 gives the standard enthalpies and entropies of reaction for the dissolution of hydrogen in different metals. With equation (11.14), they can be used to estimate the solubility of hydrogen at different temperatures.

Table 11.20 Standard enthalpies and entropies of reaction for the dissolution of hydrogen in different metals [8].

Group	Metal	Crystal class	$\Delta H°$ (KJ/mol H)	$\Delta S°$ (J/mol H K)	Temp. (°C)
IVB	Tiα	hc	– 50.6	– 53.3	350-800
	Tiβ	cc	– 56.8	– 54.0	480-950
	Nb	cc	– 33.0	– 56.4	100-900
VB	Ta	cc	– 33.0	– 54.8	300-700
VIB	Cr	cc	52.6	– 33.6	700-1200
	Mo	cc	40.5	– 50.2	600-1100
	Feα	cc	26.3	– 50.3	300-920
	Feγ	cfc	32.4	– 42.8	900-1500
	Cu	cfc	39.4	– 46.3	400-1100
	Ni	cfc	16.2	– 49.4	200-1400
	Pd	cfc	– 9.8	– 52.8	0-300

Chemical state of dissolved hydrogen

In the crystal lattice of a metal, hydrogen normally occupies octahedral or tetrahedral interstitial positions (Figure 11.21). The octahedral sites of face centered cubic (*cfc*) lattices and of hexagonal compact (*hc*) lattices have a higher symmetry than those of body centered cubic (*bcc*) lattices, where the distance between two neighboring atoms is smaller along one axis. Interstitial sites of higher symmetry are energetically more favorable for the occupation by hydrogen, because the network is less perturbed. As a consequence, the solubility of hydrogen is slightly higher in γiron (austenite), which has a *cfc* structure than in α iron (ferrite), which has a *bcc* structure (Figure 11.22). The same behavior holds for interstitial carbon atoms.

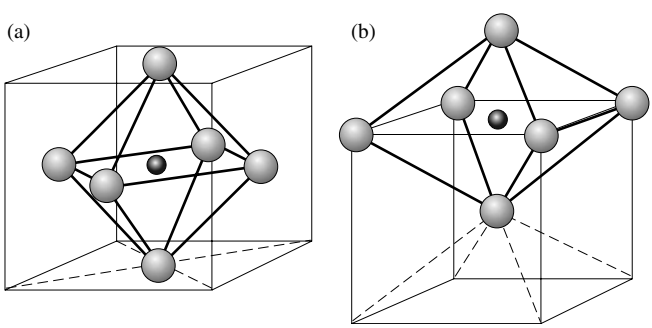


Figure 11.21 Octahedral interstitial positions in cubic face centered (cfc) and cubic body centered (bcc) crystals.

© 2007, first edition, EPFL Press

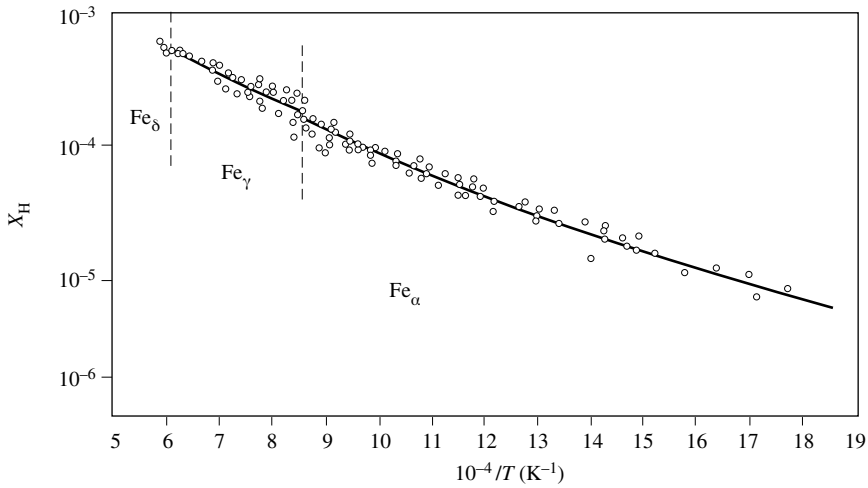


Figure 11.22 Hydrogen solubility in α, γ and δ iron as a function of temperature.

In metals, dissolved hydrogen is usually found in atomic form. Its solubility typically reaches values $X_H = 10^{-3} - 10^{-5}$ and the dissolution reaction is endothermic. According to Le Chatelier's principle, the solubility therefore increases with the temperature.

With certain metals, in particular those of groups IVB (Ti, Zr, Hf), VB (V, Nb, Ta) and Pd, the dissolved hydrogen reacts chemically, forming ***hydrides***. The atomic ratio *H*/*M* can then attain values as high as 1 or 2. The reaction being exothermic the ratio *H*/*M* decreases with increasing temperature.

Influence of defects on the solubility of hydrogen

The real solubility of hydrogen often exceeds the value calculated using thermodynamic data such as those given in Table 11.20. This discrepancy arises because metals and alloys contain defects such as vacancies, dislocations and grain boundaries which offer additional sites for hydrogen.

Figure 11.23 illustrates this phenomenon. Steel undergoes corrosion in a 0.5 M H_2SO_4 solution. The reaction gives off hydrogen, a part of which dissolves in the metal. Prior to the experiment, the steel specimens were plastically deformed to different extents by cold rolling. The graph shows that the hydrogen concentration in the steel increases linearly with the square root of the immersion time. This suggests that the rate of hydrogen absorption into the metal is controlled by diffusion. After some time, the concentration reaches a steady state value that corresponds to saturation. The saturation concentration is highest for the samples exhibiting the highest degree of deformation because cold work introduces dislocations into the metal that provide additional sites for dissolved hydrogen.

© 2007, first edition, EPFL Press

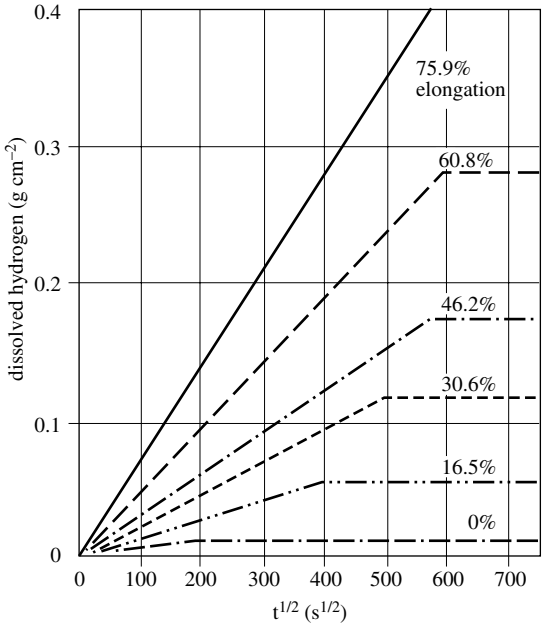


Figure 11.23 Effect of cold work on hydrogen uptake of a carbon steel: variation of dissolved hydrogen as a function of the square root of the exposure time in 0.5 M H_2SO_4 at 35 °C [10].

11.3.2 Metal damage caused by hydrogen

Hydrogen can damage a metal by chemical or physical mechanisms. Most often, hydrogen causes embrittlement of the metal and ensuing crack formation, but other forms of hydrogen damage also exist. They include:

- formation of surface hydrides;
- reaction of hydrogen with non-metallic phases;
- formation of blisters.

Cracking due to hydrogen embrittlement includes three broad categories:

- spontaneous cracking;
- cracking under the action of an applied tensile stress;
- fatigue cracking.

Superficial damage by hydride formation

When hydrogen is formed electrochemically on a metal that easily forms hydrides, a superficial layer of metal hydride may form and grow with time. Metal hydrides are generally brittle and exhibit poor mechanical resistance. The formation of a hydride layer therefore corresponds to a deterioration of the material. The described behavior, of limited practical importance, has been observed for example with titanium used as cathode in electrochemical cells.

© 2007, first edition, EPFL Press

Metal hydrides can also form at the tip of a crack if a corrosion reaction produces hydrogen there. This process is of a more dangerous nature because the presence of hydrides at the crack tip favors crack propagation by brittle fracture, as will be discussed in more detail in Section 11.4.

Reaction of hydrogen with non-metallic phases

The chemical reaction of hydrogen with non-metallic phases in an alloy deteriorates its mechanical properties. This type of reaction occurs only at elevated temperature (Chapter 9) and typically involves carbides or oxides. For example, at temperatures exceeding 200 to 300 °C, hydrogen can react with iron carbide in steels and form methane (Chapter 9). This process, referred to as *decarburizing of steel* degrades the metal in two ways. Firstly, elimination of carbides lowers the mechanical strength of the steel. Secondly, the reaction creates microscopic voids in which methane which diffuses less easily than hydrogen may become trapped. As a consequence a high pressure may build up under certain conditions and lead to formation of blisters or cracks as will be discussed below.

In a similar fashion, at elevated temperatures, hydrogen can react with oxides. For example, when copper is annealed in a hydrogen atmosphere, the following reaction can take place:

$$Cu_2O + H_{2(g)} \rightarrow 2\,Cu + H_2O_{(g)} \tag{11.15}$$

The elimination of copper oxide precipitates reduces the metal strength in the reaction zone.

Hydrogen blistering

When dissolved atomic hydrogen reacts with carbides to form methane, or if at a microscopic cavity in the metal hydrogen atoms recombine to molecules, the local gas pressure may, under certain conditions, greatly exceed the exterior pressure. Ductile metals react to this situation by plastic deformation while brittle metals form cracks. Depending on the microstructure and location of the cavity, local plastic deformation and growth of the damaged zone can lead to the formation of a blister, a phenomenon is known as ***hydrogen blistering***. Hydrogen blisters can reach diameters of several millimeters or even centimeters (Figure 11.24).

Let us consider the recombination of atomic hydrogen present in the metal to gaseous hydrogen molecules contained in a cavity:

$$2\,H_m \rightarrow H_2 \tag{11.16}$$

At equilibrium, the chemical potential of the hydrogen gas at the exterior of the metal, $\mu_{H2(ext)}$ is equal to that in the cavity $\mu_{H2(cavity)}$ and its value is twice that of the atomic hydrogen in the metal μ_{Hm}:

$$\mu_{H2(ext)} = 2\,\mu_{H_m} = \mu_{H2(cavity)} \tag{11.17}$$

© 2007, first edition, EPFL Press

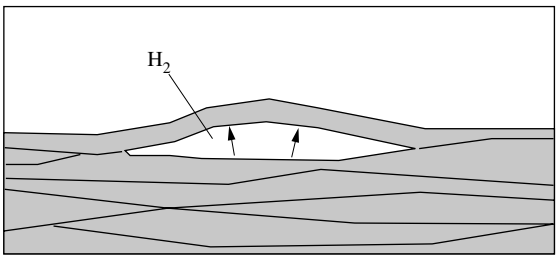


Figure 11.24 Hydrogen blistering (schematic).

The chemical potential of hydrogen in the cavity is:

$$\mu_{H2(cavity)} = \mu^0 + RT \ln P_{H2(cavity)} \tag{11.18}$$

where $P_{H2(cavity)}$ is the hydrogen pressure in the cavity. Taking into account equilibrium considerations only, the hydrogen pressure inside a cavity can not exceed the exterior pressure. However, under non-equilibrium conditions, an overpressure can develop inside the cavities.

Suppose an autoclave contains hydrogen gas at high pressure that is in equilibrium with the hydrogen contained within microscopic cavities in the wall. If the pressure in the autoclave is rapidly reduced, for example between two batch operations, without giving the system enough time to reach a new equilibrium, the pressure inside the cavities will be, for some time, much higher than the pressure acting on the wall from the outside. If this pressure difference is sufficiently important, it may cause plastic deformation of the metal separating the defect from the outside surface, resulting in the formation of a blister. Alternatively, a similar pressure difference present in a non-ductile material will lead to cracking rather than blistering.

Hydrogen blistering is not restricted to installations containing high-pressure hydrogen gas. A corrosion reaction involving the reduction of protons, for example during pickling of steel, can also be responsible for this type of damage because hydrogen evolution creates supersaturation conditions at the metal surface. During pickling of steel the oxide scales are removed from the surface by chemical dissolution in a strong acid. Simultaneously some corrosion of the metal occurs, which produces hydrogen. A part of this hydrogen diffuses into the metal and recombines at cavities. As long as the reaction continues, the pressure in the cavities will correspond to the supersaturation of hydrogen at the surface. However, when the reaction stops because the metal is being removed from the electrolyte, there will be a pressure difference between the cavity and the ambient air. In unfavorable cases, this difference can be sufficiently high to cause blister formation. To minimize corrosion and hydrogen formation, one usually adds corrosion inhibitors to the pickling acid.

Spontaneous cracking induced by hydrogen

The same mechanism that leads to blister formation in ductile materials can cause spontaneous cracking in brittle materials: atomic hydrogen dissolves in the metal, and

© 2007, first edition, EPFL Press

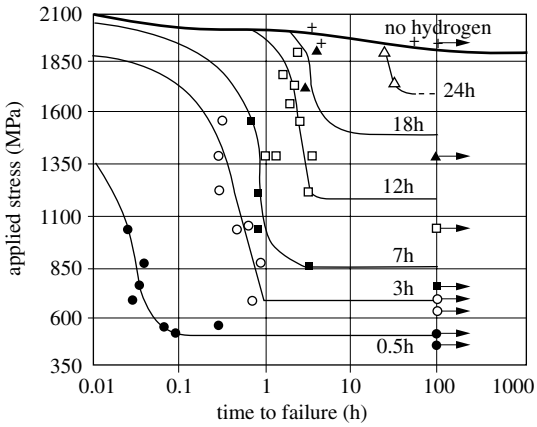


Figure 11.25 Effect of duration of heat treatment at 570 °C on time of failure of high strength steel (2070 MPa) loaded with hydrogen [11].

then, at defects recombines to form gas molecules. Under the effect of the overpressure, cracks form. The described phenomenon, often referred to as *hydrogen induced cracking* (*HIC*), shall be illustrated with the following example.

During the electrodeposition of zinc or cadmium on high strength steel, hydrogen is produced at the cathode and some of it will dissolve into the substrate. As a result, cracks may form spontaneously some time after the plated objects have been removed from the electrolyte. To avoid such delayed cracking, the plated objects must be heat treated, in order to allow the hydrogen to escape from the metal.

Figure 11.25 shows the influence of heat treatment on the time to failure of steel parts in constant load stress corrosion cracking tests. Before the tests, the specimens were saturated with hydrogen by cathodic polarization in an acid solution. They were then kept at a temperature of 570 °C for different lengths of time, so increasing amounts of hydrogen could diffuse out of the metal. The results of Figure 11.25 confirm that the thermal treatment significantly increased the time to failure. A sufficiently long heat treatment, lasting 24 hours in this case, regenerated the initial properties of the alloy. In practice, the time needed for heat treatment can be reduced by raising the temperature, because this increases the rate of diffusion.

Hydrogen induced stress cracking

The application of a tensile stress in presence of hydrogen can induce cracking and lead to premature failure of a metallic structure. Figure 11.26 illustrates this phenomenon: it shows the rate of crack propagation in a high strength carbon steel, immersed in seawater, as a function of the stress intensity factor measured under two different conditions: at open circuit and under cathodic polarization. Application of a cathodic polarization produces hydrogen, which favors cracking. This is evidenced by the fact that under cathodic polarization the value of K_{ISCC}, is lower and the crack propagation rate in the subcritical region is higher than under open circuit conditions.

© 2007, first edition, EPFL Press

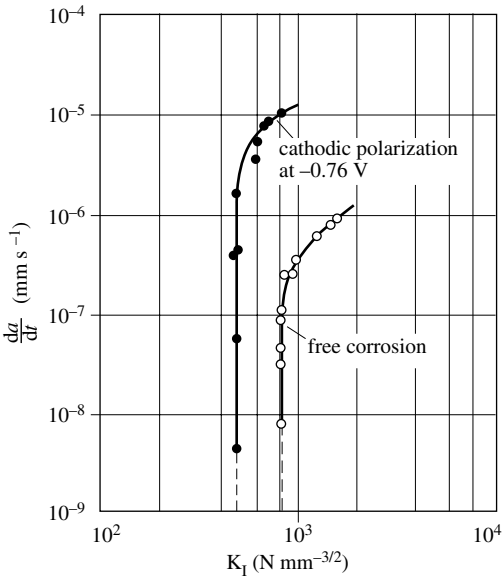


Figure 11.26 Effect of potential on crack propagation rate of high strength steel (1190 Mpa) susceptible to hydrogen embrittlement in seawater at ambient temperature [12].

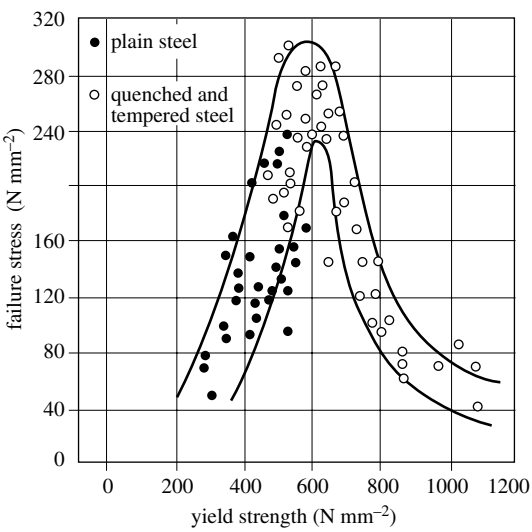


Figure 11.27 Failure stress versus yield strength for carbon steels. The failure stress was measured by placing the specimens in a 5% NaCl solution saturated with H_2S for 100 hours at ambient temperature [12].

© 2007, first edition, EPFL Press

The described behavior is sometimes referred to as ***hydrogen induced stress cracking*** (*HISC*).

Different metals are subject to hydrogen induced stress cracking, but the phenomenon is of particular importance in high strength steels ($\sigma_e > 600$ N/mm^2). In the presence of hydrogen, the ultimate tensile strength of these metals decreases drastically and may become comparable or even lower than that of ordinary carbon steel. Figure 11.27 shows the failure stress of different steels, measured in the presence of hydrogen, as a function of their yield strength. In these experiments, the failure stress has been defined as the maximum stress at with the metal resists for 100 hours in a constant load test carried out in a 5% NaCl solution at pH 3, saturated with H_2S. The presence of H_2S is known to favor hydrogen cracking. The data shown indicate that the failure stress goes through a maximum at $\sigma_e = 600$ N mm^{-2} , then strongly decreases for steels of high yield strength.

Cathodic polarization of a metal is not the only source of hydrogen that is able to induce cracking in presence of a tensile stress, either applied or residual. Gaseous hydrogen and corrosion reactions producing hydrogen have a similar effect. Of particular importance are corrosion reactions that take place at the tip of a crack where they can cause local hydrogen embrittlement of the metal, thus favoring crack propagation by brittle fracture. This type of deterioration, often referred to as ***hydrogen induced stress corrosion cracking*** *(HISCC)* , will be discussed further in Section 11.4. Many researchers do not make a distinction between *HISC* and *HISCC* because the same damage mechanisms are involved.

Fatigue cracking in presence of hydrogen

The presence of dissolved hydrogen lowers the fatigue limit of a metal subjected to cyclic loading. The phenomenon is known as ***hydrogen induced fatigue cracking***. If the hydrogen is formed by a corrosion reaction at the tip of a crack, the situation is sometimes referred to as *hydrogen induced corrosion fatigue cracking*. In both cases, the hydrogen embrittles the metal at the crack tip and thus facilitates crack propagation. These phenomena will be discussed in more detail in Section 11.5.

11.3.3 Reaction steps in hydrogen induced stress cracking

Figure 11.28 schematically represents a crack in a metal that is in contact with gaseous hydrogen. The crack propagates by successive brittle fractures, which result from the presence of hydrogen in the zone of maximum stress, just in front of the crack tip. Three reaction steps participate in the accumulation of hydrogen in this zone, and therefore determine the rate of crack propagation:

- the adsorption of hydrogen;
- the dissolution of the adsorbed hydrogen into the metal;
- the diffusion of the dissolved hydrogen towards the zone of maximum stress.

© 2007, first edition, EPFL Press

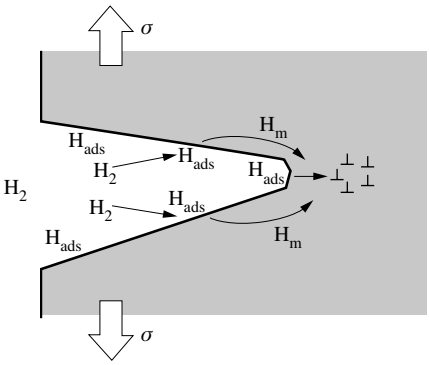


Figure 11.28 Crack propagation by hydrogen embrittlement (schematic).

Adsorption

Hydrogen adsorbs onto metallic surfaces as atoms, according to the equilibrium:

$$H_2 = 2\,H_{ads} \tag{11.19}$$

The hydrogen surface coverage, θ_H, increases with the partial pressure of gaseous hydrogen. Assuming that the adsorption follows the Langmuir isotherm (Chapter 3) we find:

$$\frac{\theta_H^2}{1-\theta_H} = b_L\,P_{H_2} \tag{11.20}$$

Here b_L represents the adsorption coefficient. From (10.20) it follows that for $\theta_H << 1$, the surface coverage is proportional to the square root of the hydrogen partial pressure:

$$\theta_H = b_L^{1/2}\,P_{H_2}^{1/2} \tag{11.21}$$

A corrosion reaction or a cathodic polarization forms adsorbed hydrogen by reduction of protons or of water molecules. The hydrogen coverage of the electrode surface under these conditions also obeys equation (11.21). However, the effective partial pressure of hydrogen that determines θ_H will usually be much higher than the ambient hydrogen pressure. Two reasons account for this observation. Firstly, the nucleation of a gas bubble at the surface of an electrode requires an over-pressure (Sect. 3.1). Secondly, for a bubble nucleus to grow, hydrogen molecules dissolved in the electrolyte must diffuse towards the bubble-electrolyte interface and this requires a hydrogen concentration gradient in the solution. As equilibrium conditions prevail at the bubble-electrolyte interface a supersaturation must exist away from the growing bubble. Depending on conditions, the hydrogen supersaturation near the metal surface can thus reach values that correspond to a hydrogen pressure of many bars.

© 2007, first edition, EPFL Press

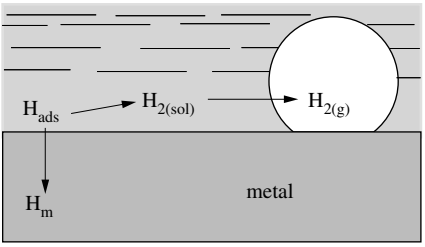


Figure 11.29 Formation of hydrogen bubble at an electrode (schematic).

Figure 11.29 schematically illustrates the described processes that lead to formation of a hydrogen bubble. It also indicates that the adsorbed hydrogen atoms formed electrochemically can react in two ways, which are in competition: they may form molecules that dissolve into the electrolyte, eventually forming a bubble, or they may diffuse into the metal as atoms. The presence of adsorbed surfactants and the surface roughness of the electrode can influence the nucleation of bubbles and, as a result, the effective hydrogen pressure.

Dissolution of hydrogen into metal

Adsorbed hydrogen dissolves more or less readily into the metal, depending on the surface state. Oxide films generally reduce the importance of hydrogen up-take by the metal, whereas adsorbed sulfides tend to facilitate it.

To study these phenomena, one measures the ***hydrogen permeability*** of a metal as a function of chosen variables using a two-compartment electrochemical cell schematically shown in Figure 11.30. The working electrode, in the form of a thin sheet, separates the two electrolyte compartments. In the cell's left-hand compartment, the electrode acts as a cathode. When a small, constant current is applied, the protons are reduced to hydrogen. A fraction of this hydrogen dissolves in the metal and diffuses across the thin sheet.

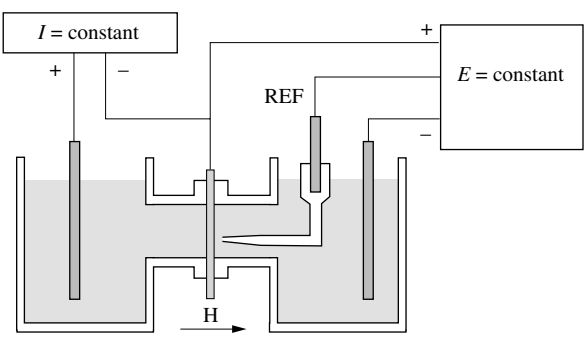


Figure 11.30 Two-compartment electrochemical cell used for the measurement of the hydrogen permeability of metals (schematic).

© 2007, first edition, EPFL Press

$$\text{Cathode compartment: } H^+ + e \rightarrow H_m \tag{11.22}$$

In the right-hand compartment, the potential of the thin sheet is set to a value that will result in the oxidation of all hydrogen that arrives at the surface by diffusion through the metal.

$$\text{Anode compartment: } H_m \rightarrow H^+ + e \tag{11.23}$$

The anodic current thus provides a direct measure of the amount of hydrogen that crosses the metal sheet at each instant. At steady state, the anodic current density $i_{a,H}$ is given by:

$$i_{a,H} = F D_H \frac{c_{H,1}}{L} \tag{11.24}$$

In this equation L represents the thickness of the metal sheet, $c_{H,1}$ designates the hydrogen concentration in the metal at the cathode surface (we assume that the concentration is zero at the anode surface, the reaction being diffusion controlled), and D_H is the diffusion coefficient for hydrogen in the metal.

By varying the electrolyte composition in the left compartment, we can study the effect of adsorbed species on the rate of hydrogen permeation. Figure 11.31 shows the results of experiments carried out with thin iron sheets (0.2 to 2 mm) in different electrolytes saturated or not with H_2S [13]. Hydrogen sulfide does not change the diffusion coefficient in the metal but it facilitates the hydrogen dissolution into iron. As a consequence, its presence in the solution increases the observed permeability. In equation (11.24), this translates into an increase of $c_{H,1}$. The described effect of H_2S

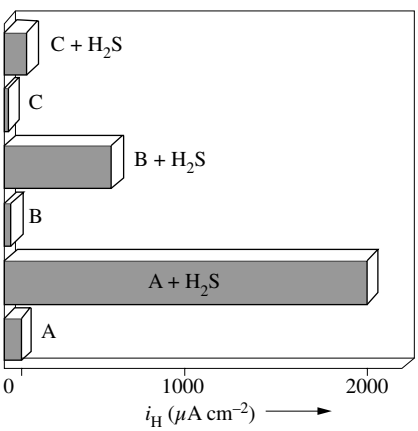


Figure 11.31 Hydrogen permeation through a Fe membrane in the presence and absence of H_2S as measured in a two-compartment cell. Electrolytes are (A) 4.7 M HCl; (B) 2.5 M H_2SO_4; (C) 2.4 M formic acid + 200 mg l^{-1} of Cl^- [13].

© 2007, first edition, EPFL Press

is of particular concern to the petroleum industry, because the saline solutions present in drilling wells usually contain sulfides.

The mechanism by which sulfides facilitate hydrogen uptake of the metal is not known with certainty. From an electrochemical point of view, the rate of hydrogen dissolution into metal should depend on the relative rates of the interfacial reactions, which are responsible for the formation of atomic hydrogen and for its recombination into molecular hydrogen. It is generally admitted (Section 5.1) that hydrogen molecules form in two steps. The first one (*Volmer reaction*) produces an adsorbed hydrogen atom.

$$H^+ + e \rightarrow H_{ads} \tag{11.25}$$

The second step leads to formation of molecular hydrogen by either the electrochemical reaction of an adsorbed hydrogen atom with a proton (*Heyrovsky reaction*) or by the chemical reaction between two adsorbed hydrogen atoms (*Tafel reaction)*:

$$H_{ads} + H^+ \rightarrow H_2 \tag{11.26}$$

$$H_{ads} + H_{ads} \rightarrow H_2 \tag{11.27}$$

Alternatively, the adsorbed hydrogen can also diffuse into the metal:

$$H_{ads} \rightarrow H_m \tag{11.28}$$

On most metals, the Volmer-Heyrovsky mechanism applies with the first step being rate limiting (Chapter 5). In this case, the adsorbed hydrogen concentration is relatively low and, far from equilibrium, does not vary much with the potential. On the other hand, if the second step is rate limiting, the activity of the adsorbed hydrogen increases with the overpotential and reaches higher values. Adsorbed species that reduce the recombination rate (Heyrovsky or Tafel reaction) relative to that of the Volmer reaction, therefore will promote the dissolution of hydrogen into the metal. This provides a plausible explanation of the effect of sulfides on hydrogen permeation. In addition sulfides inhibit the formation of passive oxide films, which generally slow down hydrogen uptake.

Diffusion of hydrogen inside the metal

In the metal the atomic hydrogen diffuses towards the zone of maximum plastic deformation, in front of the crack tip. The solubility reaches particularly elevated values at this location because of the high defect concentration (mainly dislocations). The diffusion rate depends on the difference in hydrogen concentration between the surface and the zone of maximum stress, and on the value of the diffusion coefficient. It can be studied using the two compartment electrochemical cell of Figure 11.30 by varying the thickness of the metal sheet electrode, all other parameters being kept constant.

© 2007, first edition, EPFL Press

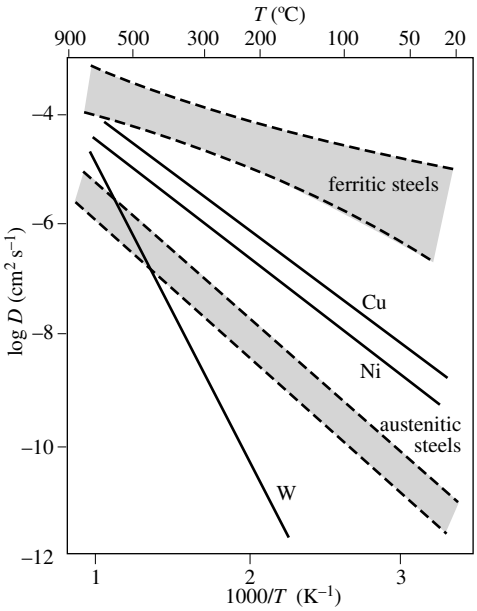


Figure 11.32 Diffusion coefficient of hydrogen as a function of temperature for different metals [14].

Experience shows that ferritic steels are more sensitive to hydrogen embrittlement than austenitic steels. And yet, the solubility of hydrogen is greater in γ iron than in α iron (Figure 11.22). The data of Figure 11.32 provide an explanation: at all temperatures the diffusion coefficient of hydrogen in α iron (*cc* structure) is significantly greater than in γ iron (*cfc* structure). In α iron, the hydrogen therefore diffuses more easily towards the zone of maximum stress, in front of the crack. However, other factors might also play a role, in particular the slip mechanism, which is different for *cfc* and *cc* metals.

Reaction mechanism of hydrogen at the crack tip

The mechanism by which the hydrogen present in the vicinity of the crack tip facilitates crack growth is not known with certainty. A plausible hypothesis is that the hydrogen renders *movement of dislocations* more difficult. As a consequence the ductility of the metal locally decreases and the zone in front of the crack tip becomes increasingly brittle. According to Figure 11.16, the local stress near a crack tip is inversely proportional to the distance r, but plastic deformation limits its value as r goes towards zero. Hydrogen embrittlement could interfere with this self-regulating mechanism, permitting the stress at the crack tip to exceed the fracture strength.

Another possibility is that the dissolved hydrogen reduces the bond energy between the metal atoms and thus the *energy of cohesion* of the metal. This may, or may not be related to hydride formation in front of the crack tip. Generally, cracks

© 2007, first edition, EPFL Press

due to hydrogen embrittlement do not propagate in a regular manner, but rather in jerks, because sequences of brittle fractures occur at the crack tip.

11.4 STRESS CORROSION CRACKING

11.4.1 Phenomenological aspects

Stress corrosion cracking (*SCC*) results from the combined action of a tensile or shear stress, applied or residual, and corrosion. Residual stresses originate either from metal shaping processes involving plastic deformation such as rolling or deep drawing or from thermal processes such as quenching or welding. Even a relatively weak residual stress, between 50 and 100% of the yield strength, sometimes suffices to cause *SCC* if a certain number of other conditions are fulfilled.

Stress corrosion cracking is normally observed only for specific metal-environment combinations and in a limited temperature interval. Table 11.33 lists a few well-known metal-environment couples that are prone to *SCC* failures. However, the number such systems grows steadily as the awareness of engineers for this phenomenon increases. The following examples serve to illustrate the specificity of the *SCC* behavior of a few alloys.

Table 11.33 Metal-environment combinations susceptible to stress corrosion cracking.

Alloy	Environment	Industries mainly concerned
Carbon steels	OH^-+H_2O $NO_3^-+H_2O$ $HCO_3^-+H_2O$	power generation
Austenitic stainless steels	Cl^-+H_2O Br^-+H_2O I^-+H_2O $OH^- +H_2O$ thiosulfate$+H_2O$	chemical and petroleum industries
Nickel alloys	H_2O	nuclear reactors
Aluminum alloys	Cl^-+H_2O Br^-+H_2O I^-+H_2O	aviation
Brass	NH_3+H_2O amines	power generation
Titanium alloys	Cl^-+H_2O, Cl^-+CH_3OH Br^-+H_2O, Br^-+CH_3OH I^-+H_2O, I^-+CH_3OH N_2O_4 anhydride	aerospace industry

Austenitic stainless steel

Austenitic stainless steels are sensitive to stress corrosion cracking in chloride-containing environments, above a temperature threshold that depends on the chloride-

© 2007, first edition, EPFL Press

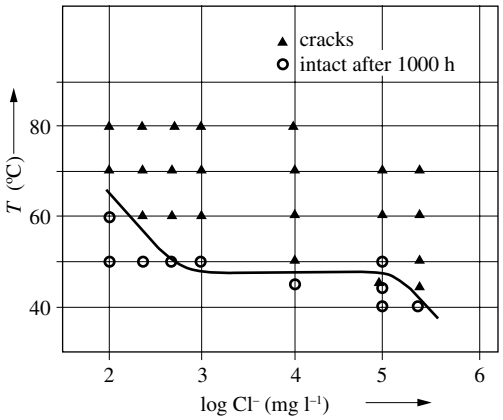


Figure 11.34 Effect of temperature and Cl^- concentration on stress corrosion cracking susceptibility of Fe18Cr-10Ni stainless steel in a $CaCl_2$ solution. Immersion time was 10,000 hours [4].

ion concentration at the metal surface. The fractures observed can be intergranular or transgranular. Figure 11.34 presents the results of *SCC* tests, carried out on U-bend specimens of a Fe-18Cr-10Ni stainless steel. The test specimens were partially immersed into $CaCl_2$ solutions of different concentrations, for a period of 10,000 hours. They were then inspected for cracks. The results indicate that stress corrosion cracking occurred if the temperature exceeded about 40 °C.

In chloride-containing environments, the sensitivity of austenitic stainless steels varies with their chromium, nickel and molybdenum content. Figure 11.35 shows the time to failure as a function of the applied load [15]. The measurements were carried out in a boiling solution of 42% $MgCl_2$. Type 304 (18-20 Cr, 8-12 Ni) resists less well than type 316, which in addition contains 2-3% molybdenum. The types 310 and

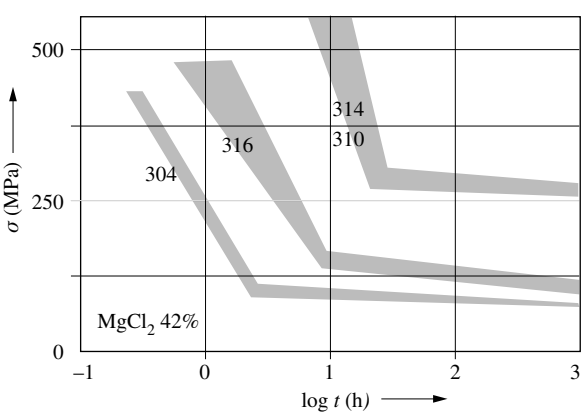


Figure 11.35 Time to failure for different austenitic stainless steels subjected to constant-load stress corrosion cracking testing in boiling magnesium chloride solution [15].

© 2007, first edition, EPFL Press

314, which contain more chromium and nickel (24-26 % Cr ,19-22 % Ni), are even more resistant.

Aluminum alloys

Certain aluminum alloys, in particular those with high mechanical strength, are sensitive to stress corrosion cracking in the presence of halogens. The sensitivity increases in the order $Cl^- < Br^- < I^-$. It varies between different alloys and depends on the thermal treatment that the metal has been subjected to.

Stress corrosion cracking of aluminum alloys is often intergranular. The chemical composition of the grain boundaries and their structure therefore play a dominant role. The SCC behavior of binary Al-Cu alloys illustrates this fact. During certain heat treatments, the intermetallic compound Al_2Cu precipitates at the grain boundaries and renders the alloy sensitive to intergranular corrosion (Chapter 7). In presence of a tensile stress this leads to cracking along the grain boundaries. The alloys Al-Mg behave similarly, because the intermetallic compounds Mg_5Al_8 and Mg_2Al_3 precipitate at the grain boundaries. A Mg content lower than 3% permits to avoid the phenomenon [16].

The thermal treatment and the concentration of the alloying elements also influence the stress-corrosion sensitivity of the ternary alloys such as Al-Cu-Mg and Al-Zn-Mg. Figure 11.36 schematically shows the effect of annealing on the mechanical strength and the stress corrosion cracking resistance of quenched alloys of this type. The intermetallic phases that precipitate preferentially at grain boundaries and grow during annealing lead first to an increase in mechanical strength with time, but the resistance to stress corrosion cracking decreases. The latter reaches a minimum at the point when the intermetallics form a coherent phase at the grain boundaries. At longer times, precipitates form mostly inside the grains and the sensitivity to intergranular *SCC* decreases. Prolonged ageing also lowers the mechanical resistance, because coalescence and growth of the precipitates diminishes their number.

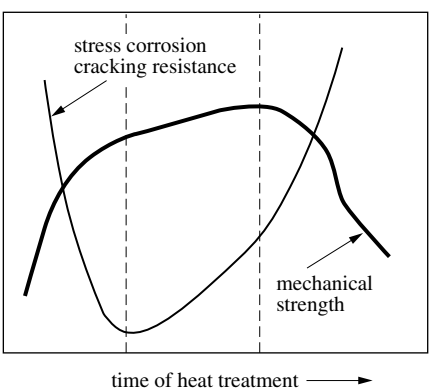


Figure 11.36 Effect of heat treatment on mechanical strength and stress corrosion cracking resistance of aluminum alloys (schematic).

© 2007, first edition, EPFL Press

Carbon steels

Carbon steels are susceptible to stress corrosion cracking in alkaline environments at elevated temperature. At the start of the industrial era, many steel-riveted boilers burst due to *SCC* because the water treatment used permitted the establishment of alkaline conditions in crevices underneath the rivets. The phenomenon was referred to as *caustic embrittlement*.

Figure 11.37 shows the influence of the potential on the time to failure for carbon steel immersed into 35% NaOH at different temperatures. The time to failure drops significantly as the temperature increases from 85 to 125 °C. For a set temperature, it exhibits a minimum at potentials between 0.7 and 0.75 V. This last observation underscores the importance of electrochemical effects in this process.

The susceptibility of carbon steel to stress corrosion cracking varies with the heat treatment, as is shown in Figure 11.38. The specimen under consideration is quenched steel with 0.25 % C, immersed in a carbonate-bicarbonate solution at 90 °C. The test was carried out at a potential of –0.41 V, which corresponds to the potential of maximum sensitivity under these conditions. The graph shows the reduction in area (equation (11.11)) measured in the corrosive environment divided by that observed in an inert environment (oil) as a function of the annealing temperature. The results demonstrate that microstructural changes resulting from heat treatment can greatly affect the sensitivity of metals to *SCC*. The observed behavior in this case has been explained by grain boundary precipitation of phases that favor intergranular corrosion [19].

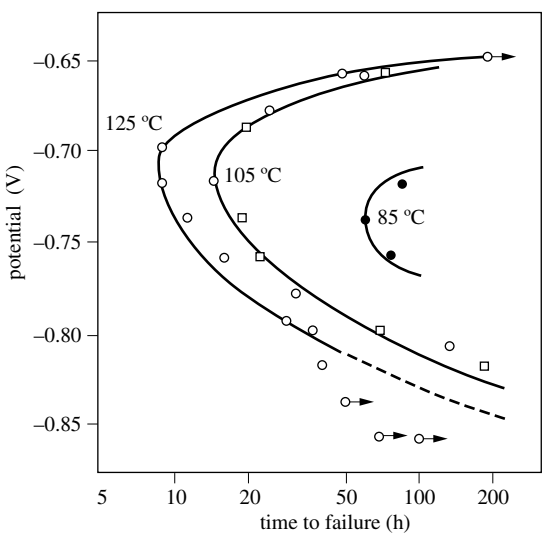


Figure 11.37 Effect of potential on time to failure of carbon steel (with 0.09% C) in a 35% NaOH solution at different temperatures [18].

© 2007, first edition, EPFL Press

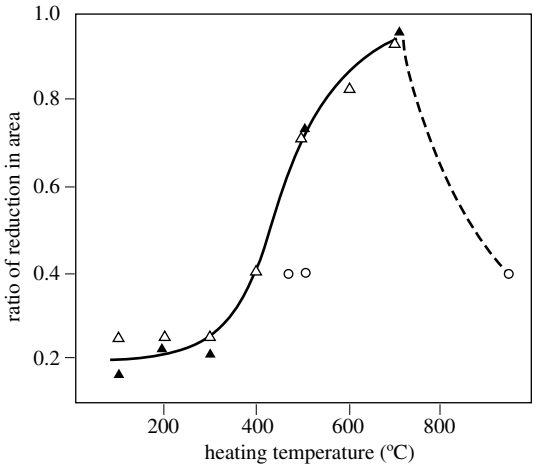


Figure 11.38 Effect of the temperature of heat treatments on hydrogen cracking susceptibility of quenched carbon steel (0.25% C) cathodically polarized at a potential of –0.14 V in carbonate-bicarbonate solution at 90 °C. Shown is the ratio of the reduction in area measured in the corrosive environment to that measured in a non-corrosive environment (oil) in SCC tests at constant deformation rate.

Nickel alloys

Intergranular stress corrosion cracking of nickel alloys and of austenitic stainless steels has sometimes been observed in the nuclear industry after prolonged exposure of these alloys to pure water, which contained trace amounts of dissolved oxygen. Although the temperature under these conditions does not exceed 100 °C, carbide precipitation at grain boundaries may occur after very long exposure times, rendering the alloys sensitive to intergranular stress corrosion cracking.

Brass

Copper-zinc alloys (brass) are extensively used in heat exchangers. They are susceptible to SCC when in contact with ammonia or other amino-compounds such as urea. The phenomenon was first observed in the British colonies when brass cartridges were stored for prolonged times near horse stables. It has since also been known as *season cracking*. Stress corrosion cracks in brasses are most often, but not always, intergranular.

11.4.2 Electrochemical considerations

The preceding discussion shows that stress corrosion cracking behavior is governed by three variables:

- the mechanical load (stress intensity at the crack tip, extent and rate of deformation);

© 2007, first edition, EPFL Press

- the material (alloy composition, microstructure, thermal treatment);
- the environment (anions, oxidizing agents, inhibitors, temperature, potential).

A complete theory of stress corrosion cracking should be able to explain the influence of all these parameters in order to predict under what conditions *SCC* is to be expected and at what rate. No such a theory exists but we can gain some useful insights by looking at the electrochemical conditions that prevail in a stress corrosion crack.

Role of surface films

Stress corrosion cracking normally occurs only with passive alloys or alloys forming a tarnish film, and it is limited to certain potential ranges. Figure 11.39 schematically shows the polarization curve of a passivating metal such as stainless steel. The shaded regions indicate the potential ranges that are favorable to stress corrosion cracking in pitting and non-pitting solutions. The potential region where hydrogen embrittlement may occur is also shown. It includes potentials that are sufficiently cathodic to allow proton reduction to take place.

In non-pitting environments stress corrosion cracking takes place preferably near the passivation potential. This applies for example to carbon steels in bicarbonate solutions and provides an explanation for the minima in the time to failure curves observed in Figure 11.37. In pitting systems, stress corrosion cracking often occurs near the pitting potential. A case in point are aluminum alloys exposed to chloride containing environments. Stress corrosion cracking in this case takes place at potentials exceeding a critical threshold that can be associated with passive film breakdown at the grain boundaries (due to the presence of precipitates). In the presence of tensile stress, the resulting selective attack at the grain boundary translates into intergranular stress corrosion cracking.

The two potential regions for *SCC* that are identified in Figure 11.39 have in common that the passive state is not stable; the metal can easily swing from active to passive behavior, and vice versa. As a consequence, the crack tip can be active, while the wall of the crack and the outside surface of the metal remain passive. Thus a corrosion cell is established where the crack tip plays the role of the anode (Figure 11.40).

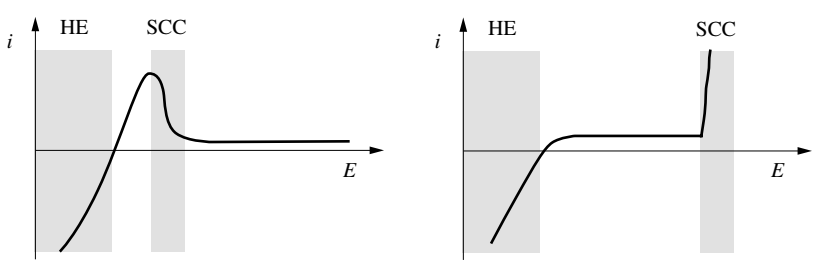


Figure 11.39 Potential domains favoring susceptibility to stress corrosion cracking (SCC) and hydrogen embrittlement (HE).

© 2007, first edition, EPFL Press

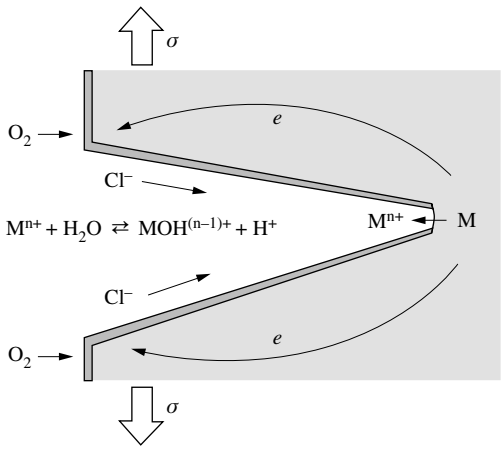


Figure 11.40 Electrochemical reactions in a stress corrosion crack.

Potential at the crack tip

The electrochemical reactions induce concentration gradients in the electrolyte present in the crack. When active dissolution takes place at the crack tip and oxygen is reduced further away or at the outside surface, the acidity of the electrolyte near the crack tip increases as a result of metal-ion hydrolysis. Electroneutrality requires that anions, such as chloride, migrate towards the crack tip to compensate the charge of the metal cations produced there. Both effects contribute to stabilize the active state of the crack tip.

The potential of an active crack tip differs from that of the passive surface away from it, but in a potentiostatic experiment, one controls the potential at the outside surface rather than that at the crack tip. The difference is all the more important as the ohmic potential drop that results from the ionic current flow in the crack is high. The described effect complicates the interpretation of electrochemical measurements in SCC experiments.

In fact, even if the applied potential is well in the passive region, the metal at the crack tip may remain in the active state. If acid conditions prevail at the crack tip the metal may even locally corrode there under simultaneous hydrogen formation.

Influence of the temperature

The schematic diagrams of Figure 11.39 allow us to interpret the influence of temperature on stress corrosion cracking. It is well known, that a higher temperature often increases the risk of stress corrosion cracking. At higher temperature passive films form less easily and their stability in pitting electrolytes is generally lower. Because of this, as the temperature increases, the corrosion potential of an alloy, which at ambient temperature would lie in the region of stable passivity, could well shift towards a potential region that is favorable for stress corrosion cracking.

© 2007, first edition, EPFL Press

Metallurgical factors also play a role. A higher temperature accelerates the rate of solid state diffusion that controls the precipitation phenomena at the grain boundaries. Under normal operating conditions the temperature of a metal will always remain well below the values applied in heat treatments. Nevertheless, slow precipitation reactions taking place during prolonged service can sensitize grain boundaries and create a danger of intergranular stress corrosion cracking. The phenomenon is favored by a higher operating temperature.

11.4.3 Mechanisms of initiation and propagation of cracks

Crack initiation

The lifetime of a structure sensitive to stress corrosion cracking is given by the time it takes to initiate a crack and the time during which the crack propagates until rupture occurs. Different phenomena, most of them difficult to control, may be at the origin of a crack nucleus and thus determine the ***crack initiation time***:

- pitting;
- intergranular corrosion;
- hydrogen accumulation in the metal due to corrosion;
- mechanical rupture of the passive film;
- microstructural defects such as inclusions, voids or microcracks.

Pits that reach a critical depth can act as crack initiation sites if they lead to a higher local stress intensity. The crack initiation time in this case corresponds to the incubation time of pits of a critical size. Alternatively, precipitation reactions at the grain boundaries can render an alloy sensitive to intergranular corrosion. The preferentially corroded grain boundary then serves as initiation site of a crack. Inclusions, preexisting microcracks, or other structural defects are also likely crack initiation sites. The crack initiation time, in this case, is defined as the time required for a crack to reach a detectable size. Crack initiation may also be the result of hydrogen formed by a corrosion reaction that may cause embrittlement of the metal or of successive ruptures of a passive film or tarnish layer, but these mechanisms are more important for the propagation than the initiation of cracks. Because of the multitude of possible crack initiation mechanisms, and because of the statistical nature of the phenomenon, it is not possible to predict the crack initiation time from first principles.

Crack propagation

Two distinct mechanisms can contribute to the growth of cracks:

- loss of material at the crack tip due to anodic dissolution;
- successive brittle fractures due to hydrogen embrittlement.

Either of these mechanisms controls the rate of crack growth and the ***crack propagation time***, which is the time a crack grows until fracture occurs.

© 2007, first edition, EPFL Press

Crack propagation by anodic dissolution

Crack propagation by anodic dissolution, implies that a corrosion cell develops between the active crack tip and the remainder of the surface, which remains passive. The crack grows by anodic dissolution at the crack tip, whereas a passive film protects the walls of the crack. The rate of crack growth therefore is proportional to the anodic partial current at the crack tip. Because the anodic surface is much smaller than the cathodic surface (in fact, one does not know exactly what is the surface area of a crack tip), the anodic partial current density at the crack tip can reach very high values on the order of several A cm^{-2}. Such high current densities correspond to substantial rates of crack growth.

Several theories have been proposed to explain the role of stress for crack growth. The widely accepted ***slip dissolution model*** assumes that plastic deformation at the crack tip occurs. The emergence of dislocations in the slip planes causes successive fractures of the passive film which lead to exposure of bare metal to the solution (Figure 11.41). The exposed microscopic areas undergo some active dissolution and then repassivate. During this short period of time, the dissolution rate locally can reach sufficiently high values to account for observed crack propagation rates. The activation-repassivation cycles repeat themselves as long as plastic deformation at the crack tip continues.

The amount of metal dissolved per cycle depends on the electrochemical parameters of the system, such as the passivation current density and the repassivation rate. On the other hand, the rate of deformation determines the frequency of activation. In Figure 11.42, crack propagation rates are plotted as a function of the repassivation

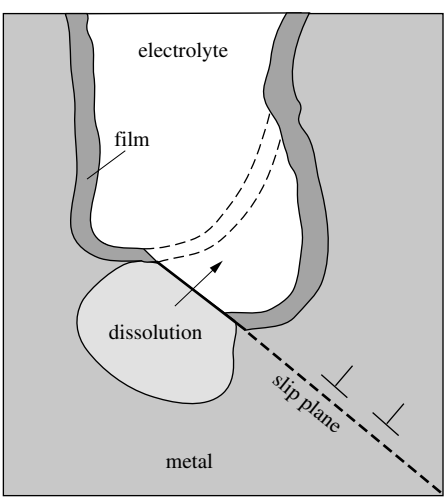


Figure 11.41 Slip-dissolution model of stress corrosion cracking. Due to slip, the metal unprotected by the passive film enters into contact with the electrolyte, thus allowing active dissolution at the crack tip for a certain period of time.

© 2007, first edition, EPFL Press

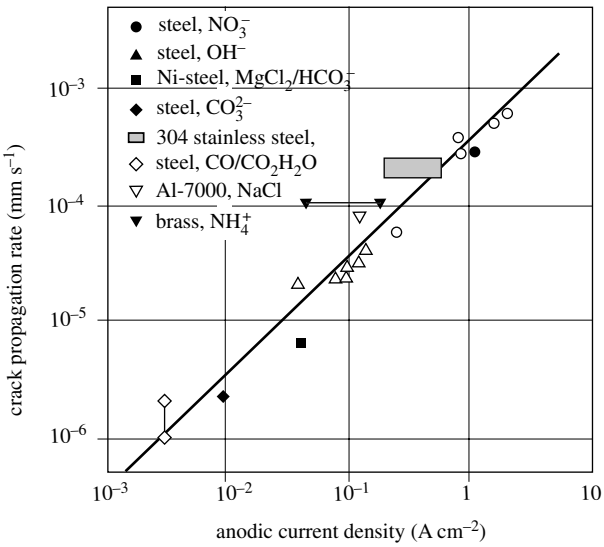


Figure 11.42 Crack propagation rate in pre-cracked specimens versus anodic current measured on flat electrodes at constant strain rate [20].

current density for various metals and environments. The anodic currents were measured on non-cracked specimens, subjected to plastic deformation, while the potential was maintained at a constant value in the passive region. The linear relationship between the rate of crack growth and the measured currents supports the described slip dissolution mechanism.

The slip dissolution model explains, at least qualitatively, a certain number of experimental observations. The propagation rate depends on a delicate equilibrium between the activation rate that is governed by the applied strain rate and the passivation rate, which only depends on electrochemical parameters. If activation is significantly more rapid than passivation, the crack tip widens because of excessive anodic dissolution. As a consequence, the stress intensity decreases, and the crack propagation comes to a stop. If, on the other hand, activation is much slower than repassivation, active dissolution does not occur and the crack propagation rate becomes negligibly small. The cracks therefore can grow only under specific conditions with regard to the rate of metal deformation and of applied potential. They correspond to situations where the rates of activation and of passivation are comparable. Only metal-environment systems that satisfy the described conditions are sensitive to stress corrosion cracking.

The slip dissolution model assumes that plastic deformation at the crack tip is responsible for the activation. But other mechanisms can have the same effect. Tensile stress at the crack tip could, for example, break a brittle tarnish film or passive oxide film, thereby exposing the base metal to the electrolyte. Selective dissolution of alloy components at the crack tip could locally weaken the metal matrix and thus permit

© 2007, first edition, EPFL Press

crack propagation by successive brittle fractures [21,22]. Another model proposes that vacancies formed during anodic dissolution coalesce at the crack tip and thus account for crack growth [23]. Alternatively, it is well known that the precipitation of certain phases such as carbides etc. facilitates intergranular corrosion. Application of a tensile stress then permits intergranular crack growth by continuously exposing new surface to the electrolyte. The phenomenon is sometimes referred to as *stress assisted intergranular corrosion* is one of several mechanisms of stress corrosion cracking.

Propagation by successive brittle fracture

Numerous observations show that cracks often advance by a succession of rapid jerks, which is contrary to what one would expect if crack propagation is due to anodic dissolution at the tip. In addition, the transient propagation rates can be very fast, difficult to reconcile with anodic current densities at the crack tip even if these can be quite high as discussed above. Other mechanisms therefore contribute to crack propagation in these cases, a most likely one is crack propagation by successive brittle fractures of the metal. A case in point is ***hydrogen induced stress corrosion cracking*** (*HISCC*).

Let us assume that the plastic deformation at the crack tip activates the metal, as in the previous model. If we are dealing with a non-noble metal and if the corrosion potential at the crack tip is relatively low, a corrosion reaction can take place that produces hydrogen by reduction of protons. The hydrogen that locally dissolves into the metal leads to embrittlement in the zone of maximum stress in front of the crack tip.

Cracking occurs periodically in an abrupt manner, whenever the metal has accumulated a critical quantity of hydrogen in the zone of maximum stress. According to this model, cracks propagate by brittle fracture, whereas corrosion provides a local source of hydrogen and thus determines to a large extent the average rate of crack growth. The model of crack propagation by brittle fracture thus can explain the high instantaneous rates of crack growth and their discontinuous nature and it is consistent with the observed surface morphology after rupture, which corresponds to a brittle fracture. Depending on the conditions, ruptures may be transgranular or intergranular. Crack growth by successive brittle fractures is probably responsible for stress corrosion cracking of high strength steels and of certain aluminum alloys. It may also apply to *SCC* of titanium alloys, where formation of hydrides may possibly enhance the local embrittlement of the metal.

Other mechanisms can also contribute to crack propagation by brittle fracture. In particular, ***dealloying*** by anodic dissolution occurring at the crack tip could locally lower the mechanical strength of the metal and render it brittle. The ensuing brittle fractures would then allow the crack to advance step by step. According to this mechanism, which has been proposed for copper alloys and for others, selective anodic dissolution serves to weaken the alloy at the crack tip, whereas the crack propagation is governed by mechanical processes [21,22].

© 2007, first edition, EPFL Press

11.5 CORROSION FATIGUE

11.5.1 Experimental study of corrosion fatigue

For the experimental study of corrosion fatigue, one applies a periodic tensile stress or a periodic strain to the metal by means of a fatigue testing machine. During testing the specimens are exposed to a corrosive environment, either gaseous or liquid. For tests carried out in electrolyte solution, it is possible to vary the potential of the metal with respect to a reference electrode, using a potentiostat. In order to evaluate the influence of corrosion, one usually compares results obtained under corrosive and non-corrosive conditions. In the following we distinguish two kinds of corrosion fatigue tests, defined by the type of specimen:

- tests using non-precracked specimens;
- tests using precracked specimens.

Corrosion fatigue testing with non-precracked specimens

The specimens are either smooth or notched. They are subjected to either a cyclic load or a cyclic deformation. The cyclic parameter may follow a sinusoidal, triangular, or trapezoidal function or correspond to some other waveform. Figure 11.43 defines the three mechanical parameters of a corrosion fatigue test with a sinusoidal applied load:

- maximum stress, σ_{max};
- minimum stress, σ_{min} ;
- cycle time, t_{cycle}.

A positive sign is assigned to tensile stress. The figure shows two other parameters that are often used: the average stress, $\sigma_m = (\sigma_{max} + \sigma_{min})/2$ and the maximum stress variation, $\Delta\sigma = \sigma_{max} - \sigma_{min}$. The stress ratio $R = \sigma_{max}/\sigma_{min}$ is also useful for the characterization of the mechanical loading.

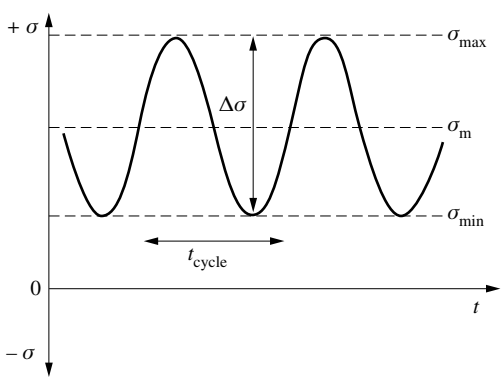


Figure 11.43 Different parameters characterizing the fatigue corrosion test.

© 2007, first edition, EPFL Press

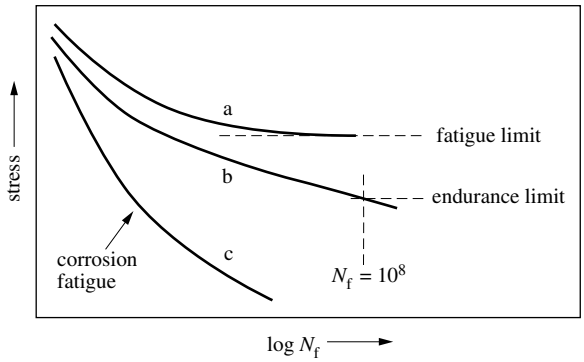


Figure 11.44 Wöhler curve in inert and in corrosive environment (schematic). The figure shows (a) a metal with a fatigue limit; (b) a metal without a fatigue limit; and (c) a metal undergoing corrosion fatigue.

In fatigue tests with non-precracked specimens one measures the ***time to failure*** t_f, which is usually expressed in number of cycles:

$$N_f = \nu\, t_f \tag{11.29}$$

Here, N_f designates the number of cycles to failure and ν is the frequency of the applied load. The graphical representation of the results of fatigue testing in the $\Delta\sigma$-$\log N_f$ plane is called the *Wöhler curve*. It allows us to determine the ***fatigue limit***, defined as the stress below which no fatigue failures occur (Figure 11.44). Certain metals do not exhibit a well defined fatigue limit. In this case, one defines the ***endurance limit*** as the maximum stress at which the metal specimen resists for at least 10^8 cycles.

Corrosion reduces the time to failure and eliminates the fatigue limit. Its detrimental effect is particularly important in ***low cycle fatigue***. In this kind of fatigue, the applied stress is sufficiently high to cause plastic deformation and as a consequence the time to failure, even in the absence of corrosion, is usually less than 10^4 cycles.

Factors affecting the results obtained with non-precracked specimens

The time to failure observed in corrosion fatigue testing with non-precracked specimens not only depends on the applied mechanical load and the corrosivity of the environment but also on other variables such as:

- the surface state of the metal;
- the shape of the specimen;
- the type of mechanical loading (tensile or bending).

The surface state of the metal influences the crack initiation time. On rougher surfaces, cracks nucleate more easily than on a smooth surface. Surface stresses also play a role: a surface under compression, for example following sand blasting or shot peening, generally exhibits a higher crack initiation time than an annealed surface.

© 2007, first edition, EPFL Press

The results of corrosion fatigue experiments with non-precracked specimens depend on the shape and size of the specimens used. Normally these parameters are defined in standard practices and norms. The presence of a notch permits to localize the fracture and reduces the time to failure.

In practice, different types of fatigue testing machines are used for corrosion fatigue experiments, using either traction or bending. The time to failure varies between different experimental arrangements. A direct comparison of test results obtained with non-precracked specimens is possible only if the conditions with respect to all of the discussed variables are the same.

The time to failure measured on non-precracked specimens represents the sum of the initiation and propagation times of a crack. Unfortunately, crack initiation, which often is a slow step, is a random phenomenon. As a result, the reproducibility of corrosion fatigue tests carried out with non-precracked specimens is usually mediocre. Furthermore, to determine a Wöhler curve, a large number of specimens is needed. Statistical methods are usually applied to determine the endurance limit in corrosion fatigue tests.

Corrosion fatigue testing with precracked specimens

In this type of experiment one measures the crack propagation rate as a function of the variation ΔK_I of the stress intensity factor. The results are interpreted in terms of fracture mechanics in the same manner as in stress corrosion cracking. Corrosive

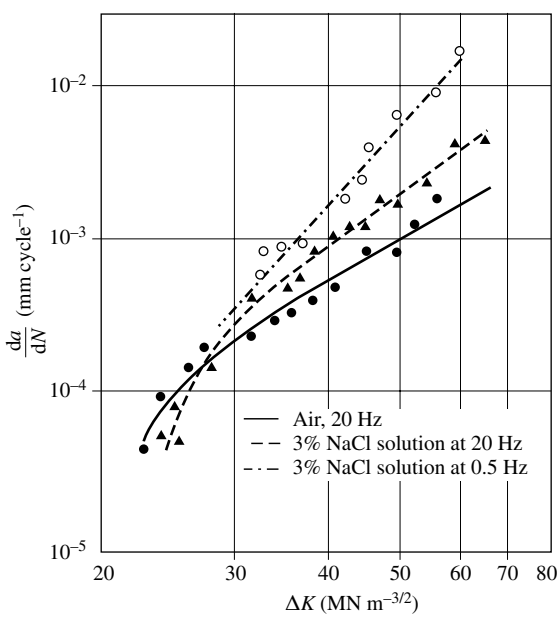


Figure 11.45 Effect of environment and applied frequency on crack propagation rate during corrosion-fatique testing of a duplex stainless steel [24].

© 2007, first edition, EPFL Press

environments accelerate crack propagation for any given value of ΔK_I, as is illustrated by the data of Figure 11.45. The figure shows the crack propagation rate, da/dN, as a function of ΔK_I for a duplex stainless steel exposed to air or immersed in a solution of 3% NaCl. Above a certain stress intensity, corresponding to approximately 30 MN m$^{-3/2}$, the results approximately follow a relation of the type:

$$\frac{\mathrm{d}a}{\mathrm{d}N} = k\,\Delta K_I^{\mathrm{m}} \tag{11.30}$$

In this equation, k and m are constants obtained by fitting the experimental data. At a frequency of 20 Hz, the crack growth rate in the NaCl solution is higher than in air indicating that corrosion fatigue occurred. The crack growth rate in the NaCl solution also depends on frequency: it is higher at 0.5 Hz than at 20 Hz. Indeed, in corrosion fatigue, one often observes that the crack propagation rate is higher at lower frequencies because a low frequency allows more time for corrosion to act at the crack tip.

The results of experiments based on fracture mechanics are, in principle, independent of the particular specimen geometry. They therefore can be compared even if obtained with different experimental arrangements. On the other hand, precracked specimens do not allow for the study of crack initiation time, which, in practice, is often a determining factor for the lifetime of a structure.

11.5.2 Mechanism of corrosion fatigue

Comparison of corrosion fatigue with stress corrosion cracking

Corrosion-fatigue fractures, just as those due to stress corrosion cracking, can be either intergranular or transgranular and in general they exhibit a morphology typical for brittle fractures. On the other hand, the two phenomena differ in a number of ways.

In contrast to stress corrosion cracking, corrosion fatigue is not just limited to specific metal-environment systems, but affects all metals that are subjected to cyclic tensile stresses in a corrosive environment. Figure 11.46 illustrates the non-specificity of corrosion fatigue [24]. The number of cycles to failure is plotted as a function of the amplitude of the applied stress for duplex stainless steel samples exposed to air, or immersed in a solution of either 3% NaCl or 0.05M H_2SO_4. The specimens degrade more rapidly in sulfuric acid in spite of the fact that this solution does not cause stress corrosion cracking for this type of steel, contrary to the NaCl solution.

In general, the importance of corrosion fatigue in a given environment increases with the intrinsic corrosion rate of the metal. In Figure 11.46, for example, the rate of uniform corrosion of the stainless steel is more important in sulfuric acid than in neutral NaCl solution. The time to failure for the specimen in sulfuric acid is therefore shorter. The results of Figure 11.47 confirm the importance of general corrosion [25]. Corrosion-fatigue tests were carried out on non-precracked specimens of carbon steel in a solution of 3% NaCl. Different anodic current densities were applied in order to

© 2007, first edition, EPFL Press

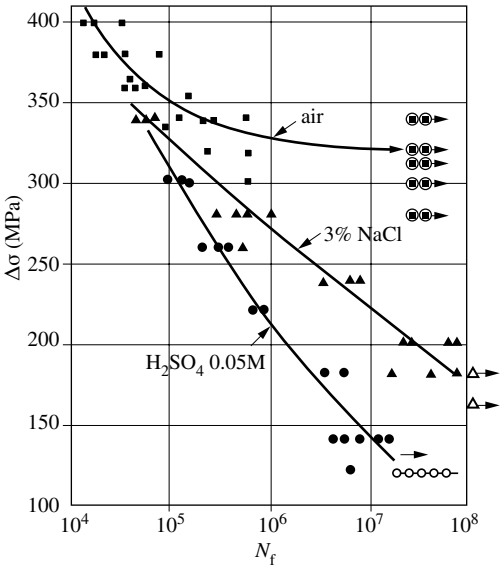


Figure 11.46 Effect of corrosive environment on number of cycles to failure of Fe-17Cr-12Ni stainless steel [24].

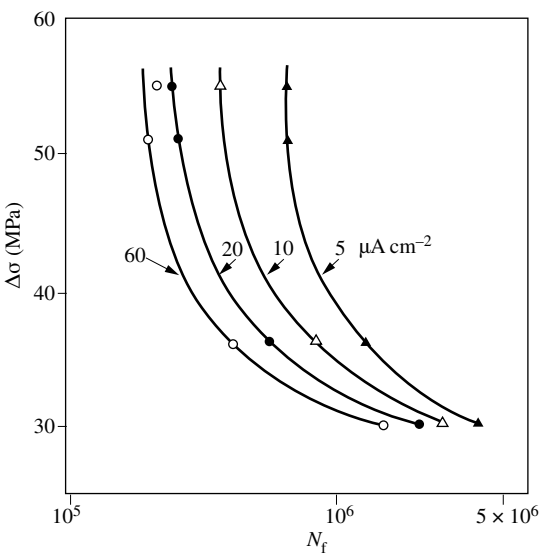


Figure 11.47 Effect of anodic current density on cycles to failure of a type AISI 4140 steel during corrosion fatigue testing in 3% NaCl at ambient temperature [25].

© 2007, first edition, EPFL Press

simulate the corrosion rate of the metal. The data indicate that the time to failure becomes shorter as the anodic current density is increased.

Crack initiation

The mechanisms that control crack initiation in corrosion fatigue are poorly understood. It is generally believed, however, that corrosion in its various forms (pitting, intergranular, acid, at inclusions etc.) plays a dominant role. The observed effect of the general corrosivity of the environment on the time to failure in Figure 11.46 supports this statement, because the results were obtained on non-precracked specimens and this kind of test provides mostly an indication of the crack initiation time.

Crack propagation

Three distinct phenomena can contribute to crack propagation:

- mechanical fatigue;
- corrosion fatigue;
- stress corrosion cracking.

Figure 11.48 shows schematically the influence of the three mechanisms on the shape of the curves that represent the crack propagation rate as a function of the stress intensity. In Figure 11.48(a), corrosion fatigue and mechanical fatigue are compared. In the subcritical region, the cracks propagate more rapidly in a corrosive environment. Equation (11.30) can be used to describe the behavior corresponding to this "true" corrosion fatigue.

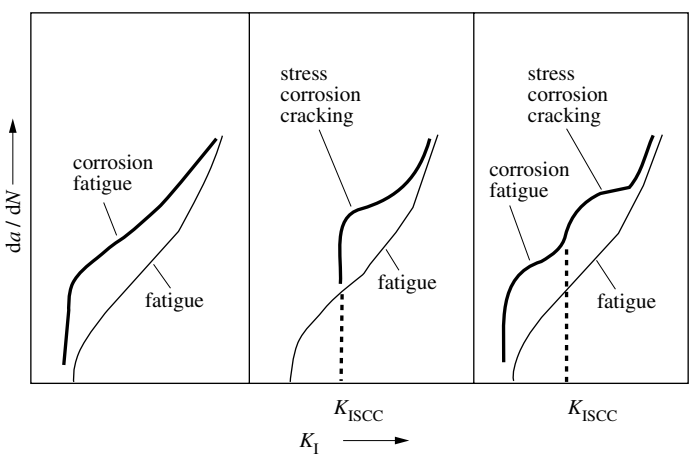


Figure 11.48 Enhancement of fatigue crack propagation rate by corrosion (schematic, logarithmic scale): (a) corrosion fatigue, (b) stress corrosion cracking, (c) a combination of corrosion fatigue and stress corrosion cracking.

© 2007, first edition, EPFL Press

Part (b) of the Figure 11.48 shows the superposition of the behaviors of mechanical fatigue and stress corrosion cracking. In general, the influence of the latter shows up mostly at low strain rates ($d\varepsilon/dy = 10^{-3} - 10^{-7}\ s^{-1}$). In the graph that represents da/dN as a function of log K_I, a step appears at the point where the stress intensity factor that corresponds to the maximum value of the applied cyclic stress, reaches the value K_{ISCC}, the threshold stress intensity for stress corrosion cracking.

Under certain conditions, the three mechanisms contribute simultaneously to crack propagation, a situation corresponding to the Figure 11.48(c), which shows a curve that exhibits two steps. Corrosion fatigue is acting even at stress intensity factors below K_{ISCC}. By analogy to K_{ISCC}, we can define a ***threshold stress intensity factor for corrosion fatigue***, K_{ICF}. In general, $K_{ICF} < K_{ISCC}$.

Figure 11.49 shows the contribution of the three crack propagation mechanisms to fatigue cracking of a high strength steel exposed to different environments [26]. In air, the crack growth is the result of mechanical fatigue. In seawater, it is caused by corrosion fatigue. In seawater containing H_2S, produced by microorganisms, a second step leading to a plateau in the da/dN curve can be identified. It indicates that stress corrosion cracking contributes to crack propagation. Under the conditions of the experiments, local embrittlement of the metal at the crack tip by hydrogen produced by corrosion is the likely cause for the contribution of *SCC* to crack growth.

Most metal-environment systems do not yield experimental crack growth curves with such well pronounced characteristics. The identification of the different mechanisms contributing to crack growth is then more difficult. One may expect that similar mechanisms as those responsible for crack growth under static conditions

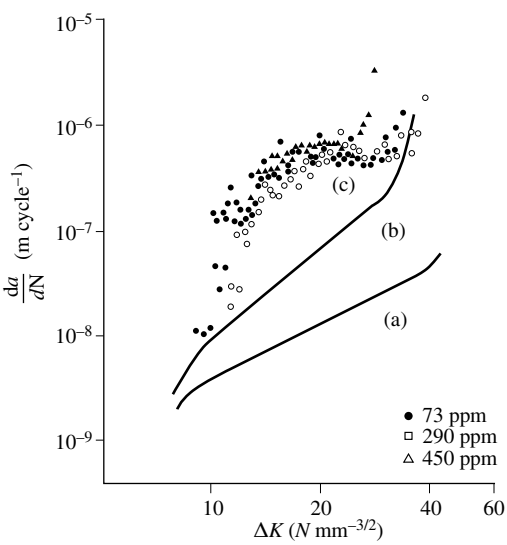


Figure 11.49 Crack propagation rate of steel (RTQ 501) subjected to fatigue testing in different environments: (a) air; (b) seawater; (c) seawater containing different concentrations of H_2S [26].

© 2007, first edition, EPFL Press

govern crack propagation during cyclic loading: intergranular corrosion, slip dissolution, hydrogen embrittlement, etc. In practice, corrosion fatigue seems to be most often governed by anodic dissolution at the crack tip. This could explain to some extent why corrosion fatigue is not limited to certain metal-environment systems, but can occur in any corrosive environment.

11.5.3 Overview of environment induced cracking phenomena

The terminology used in scientific and technical literature for the designation of environment induced cracking phenomena is not rigorous and sometimes confusing. To a large extent, this is due to the fact that the underlying crack initiation and growth mechanisms are insufficiently understood. Even for comparable mechanical conditions, these mechanisms vary between different alloys that differ in their electrochemical behavior in a given environment. In this chapter, we described three broad types of *EIC* phenomena in metals: hydrogen induced cracking, stress corrosion cracking and corrosion fatigue. They are schematically presented in Figure 11.50, grouped according to the dominating cracking mechanism and the nature of the applied stress. Three different situations are distinguished: zero applied stress, $\sigma = 0$; static tensile stress, σ = constant; and cyclic stress, $\sigma = f(\omega t)$. Two limiting crack growth mechanism are identified: hydrogen embrittlement and anodic dissolution at the crack tip. The figure highlights the fact that depending on the metal-environment system considered, both *SCC* and corrosion fatigue involve crack growth either by brittle fractures due to hydrogen or by anodic dissolution and corrosion at the crack tip. The arrows bring to mind that anodic mechanisms tend to favor intergranular cracking, whereas hydrogen embrittlement produces more likely transgranular cracks.

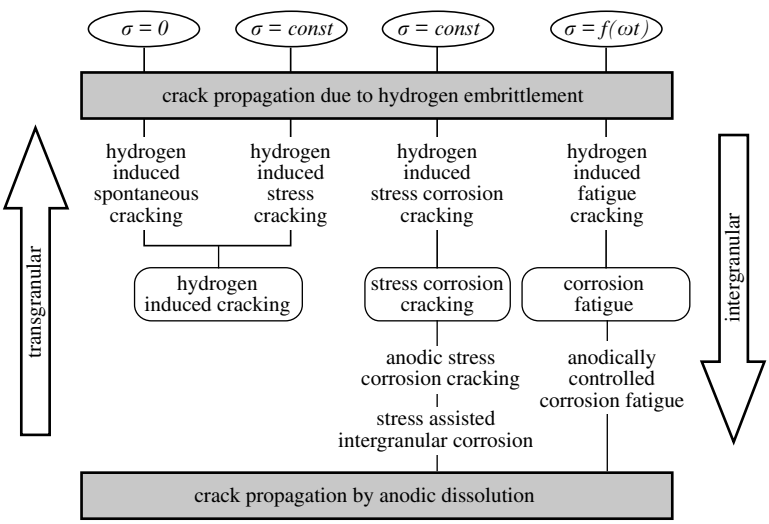


Figure 11.50 Overview of stress corrosion cracking phenomena.

© 2007, first edition, EPFL Press

Hydrogen-induced cracking includes spontaneous cracking and cracking under an applied tensile stress. The latter differs from hydrogen induced stress corrosion cracking by the source of the hydrogen, rather than by the cracking mechanism. Stress corrosion cracking involving anodic dissolution, in many cases, resembles intergranular corrosion, the tensile stress facilitating the growth of cracks along the grain boundaries. There may be intermediate cases of stress corrosion cracking or corrosion fatigue where crack growth involves both anodic dissolution and brittle fracture, for example if cracks grow by dealloying. In the simplified scheme presented in Figure 11.50, such phenomena are not included.

General Bibliography

R. W. Staehle, J. Hochmann, R. D. McCright, J. E. Slater, eds., *Stress Corrosion Cracking and Hydrogen Embrittlement of Iron Base Alloys,* NACE-5, Natl. Assoc. of Corrosion Engineers, Houston Texas (1977) 1210 pp.

D. O. Sprowls, *Metals Handbook,* 9th edition, Vol. 13, ASM Metals, Park Ohio (1987) pp. 245-282.

R. C. Newman, *Stress Corrosion Cracking Mechanisms, Corrosion mechanisms in theory and Practice* 2nd edition, P. Marcus, ed., Marcel Dekker New York 2002, pp. 399-450.

References

[1] M. H. Kamdar, *Metals Handbook,* 9th edition, Vol. 13, ASM Metals, Park Ohio (1987) pp. 171-184.

[2] Photographs generously supplied by BASF, Ludwigshafen.

[3] Norm ISO No7539 (1989) 7 pp.

[4] G. Herbsleb, B. Pfeiffer, *Werkstoffe und Korrosion* 35, 254 (1984).

[5] M. O. Speidel, *Corrosion in Power Generating Equipment*, M. O. Speidel, A. Atrens, eds., Plenum Press, New York (1984) p. 85.

[6] J. R. Scully, P. J. Moran, *Corrosion,* NACE 44, 176 (1987).

[7] H. Margot-Marette, G. Bardou, J. C. Charbonnier, *Corr. Sci.* 27, 1009 (1987).

[8] P. T. Gallagher, W. A. Oates, *Trans. AIME* 245, 179 (1969).

[9] J. R. G. da Silva, S. W. Stafford, R. B. McLellan, *J. Less Common Metals* 49, 407 (1976).

[10] J. Weber, *Oberfläche/Surface* 28, No. 4, 8 (1987).

[11] B. Craig, *Metals Handbook,* 9th edition, Vol. 13, ASM Metals, Park Ohio 1987, p. 163.

[12] H. Pircher, *Wasserstoff und Korrosion*, D. Kuron, ed., Verlag Irene Kuron, Leverkusen (1986) pp. 75-124.

[13] G. Schmidt, *Wasserstoff und Korrosion*, D. Kuron, ed., Verlag Irene Kuron, Leverkusen (1986) pp. 255-299.

[14] P. Tison, R. Broudeur, J. P. Fidelle, B. Hocheid, *Proc. Second International Congress on Hydrogen in Metals*, Pergamon Press, Oxford (1978) pp. 1-8.

© 2007, first edition, EPFL Press

[15] D. O. Sprowls, *Metals Handbook*, 9th edition, Vol. 13, ASM Metals, Park Ohio (1987) p. 272.
[16] J. C. Scully, *Corrosion*, 2nd edition, Vol. I, H. H. Shreir, ed., Butterworth, London (1976) pp. 8-91.
[17] M. O. Speidel, M. V. Hyatt, *Advances in Corrosion Science and Technology*, M. G. Fontana, R. Staehle, eds., Vol. 2, Plenum Press, New York (1972) pp. 115-335, p.255.
[18] H. H. Uhlig, R. W. Revie, *Corrosion and Corrosion Control*, 3rd edition, John Wiley, New York (1985) p. 138.
[19] R. N. Parkins, *Corrosion in Power Generating Equipment*, M. O. Speidel, A. Atrens eds., Plenum Press, New York (1984) pp. 53-82.
[20] R. N. Parkins, *Corrosion*, 2nd edition, Vol. I, H. H. Shreir, ed., Butterworth, London (1976) pp. 8-18.
[21] R. C. Newman, *Corrosion Mechanisms in Theory and Practice,* 2nd edition, P. Marcus, ed., Marcel Dekker, New York (2002) p. 399.
[22] K. Sieradzki, J. S. Kim, A. T. Cole, R. C. Newman, *J. Electrochem. Soc.,* 134, 1635 (1987).
[23] J. R. Galvele, *Corrosion* 55, 723 (1999).
[24] P. Combrade, *Les aciers inoxydables*, P. Lacombe, B. Baroux, G. Béranger, eds., Les éditions de physiques, Paris (1990) p. 339.
[25] H. H. Uhlig, *Corrosion Fatigue* NACE-2, A. J. Mc Elivy, R. Staehle, eds., Natl. Assoc. Corrosion Engineers, Houston Texas, (1972) p. 270.
[26] C. J. Thomas, R. G. J. Edyvean, R. Brook, I. M. Austen, *Corr. Sci*. 27, 1197 (1987).

© 2007, first edition, EPFL Press

CHAPTER 12

PROTECTION OF ENGINEERING SYSTEMS AGAINST CORROSION

12.1 PREVENTIVE MEASURES

12.1.1 Introduction

The fight against corrosion must begin in the planning stage of a project. In other words, it is necessary to take corrosion into account from the very beginning of a project and during every stage to its completion. The goal is to guarantee a predetermined lifetime for an installation, machine or device at minimum cost, taking into account both the investment and the maintenance costs. Furthermore, the chosen solution must be compatible with environmental regulations, permitting the recycling or elimination of the different components at the end of their use. In many applications (aerospace, nuclear reactors, chemical plants, etc.), safety of operation outweighs all other considerations.

A number of different approaches and methods can be used to minimize the corrosion of engineering structures:

- appropriate design;
- judicious choice of materials;
- protective coatings;
- corrosion inhibitors;
- electrochemical protection.

The first two are preventive measures. Taken at the right time, they allow engineers to minimize corrosion problems at marginal additional cost. Protective measures such as the use of coatings, of inhibitors or of electrochemical protection usually require more specific knowledge and often need the intervention of a specialist. They usually generate costs related to maintenance and control. The distinction between preventive and protective measures is not sharp, however; depending on the application, the use of protective coatings can be thought of as a preventive or a protective measure.

The concept of ***corrosion system*** is of particular importance for the selection of methods of corrosion prevention and protection. Indeed, the corrosion behavior of a metal is influenced by a multitude of electrochemical, mechanical and physical

© 2007, first edition, EPFL Press

interactions with the environment such as corrosion cells, temperature and humidity cycles, stress, fluid flow, erosion etc. Such phenomena can have a profound effect on the life time and safety of engineering structures and must be taken into account throughout the different stages of a project. Regular maintenance and corrosion monitoring are other important aspects of the fight against corrosion.

12.1.2 Corrosion prevention by design

Proper design of engineering structures greatly reduces the risks of corrosion. Indeed, the shape of an object can critically affect its lifetime in a given corrosive environment. This general statement shall be illustrated with a few simple examples.

Humidity

Atmospheric corrosion requires the presence of humidity to form an electrolyte (Chapter 8). Preventing accumulation of humidity in certain spots by clever design therefore greatly reduces the risk of atmospheric corrosion. This aspect is particularly important for structures in civil engineering and architecture that are periodically exposed to rainwater. In a similar fashion, in the chemical industry, it is important to be able to completely empty containers when they are not in use. Figure 12.1 shows how the design of a reactor can facilitate its efficient drainage.

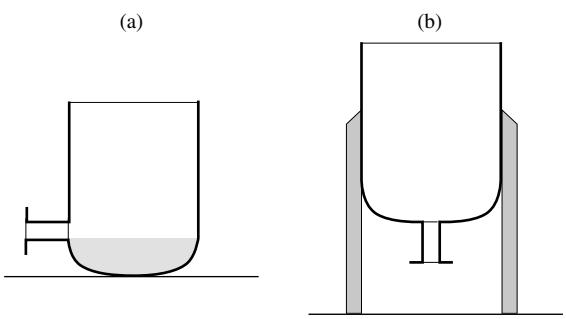


Figure 12.1 Reduction of risk of corrosion in the interior of a chemical reactor when not in use: (a) incomplete drainage which can lead to corrosion; and (b) complete drainage which reduces the corrosion risk.

In certain cases, good aeration is enough to avoid the formation of humid zones. For example, the corrosion of car doors usually develops from the inside, where the moisture accumulates, and moves outwards. Well-designed aeration vents prevent the accumulation of humidity and thus greatly decrease the rate of corrosion.

Stress concentration

The stress distribution in a given object influences its resistance towards stress corrosion cracking (Chapter 11). The tensile stress at a given spot depends critically on the shape of the object. Figure 12.2 shows how, by use of a rounded shape instead

© 2007, first edition, EPFL Press

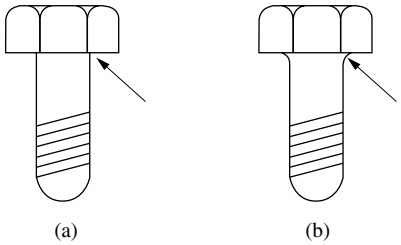


Figure 12.2 Reduction of the risk of stress corrosion cracking of a bolt: (a) high local stress; and (b) reduced local stress.

of a sharp angle, one can avoid an excessive stress concentration near the head of a bolt. Mechanical machining or thermal fabrication processes can be a source of residual stress in metallic objects, independent of the their shape. To reduce residual stresses one applies an adequate thermal treatment.

Fluid flow

A sound design of fluid flow systems reduces the risk of flow accelerated corrosion (Chapter 10). Most important is to avoid sudden changes in flow direction at sharp bends. As a rule of thumb, the radius of a pipe bend carrying an aqueous fluid should be greater or equal to three times the pipe diameter (Figure 12.3).

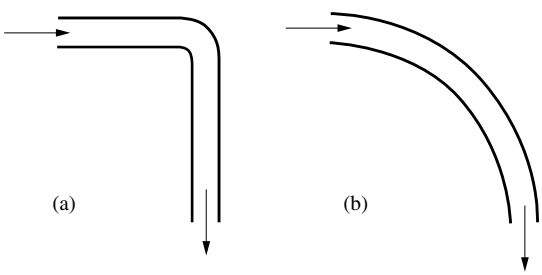


Figure 12.3 Reduction of the risk of flow accelerated corrosion at an elbow: the risk is greater for the configuration (a) than for that of (b).

Crevices

Crevices favor the formation of aeration cells (Chap. 7), and lead to crevice corrosion. Crevices in metallic structures can be largely avoided by proper design. Figure 7.13 gives an example of a crevice formed in an improperly designed pipe flange. Riveted or screw heads often form a crevice with the underlying metal. Welded or glued joints are, from this point of view, preferable to riveted or screwed assemblies.

© 2007, first edition, EPFL Press

Contact between different metals

If both an electric and an electrolytic contact are established between different metals, a galvanic cell forms that results in an accelerated corrosion of the metal with the less noble corrosion potential (Chapter 7). With an adequate selection of materials, or by placing an insulating material between the metals, this kind of corrosion can be avoided (Figure 12.4(a)). For small objects, such as electronic components, one can prevent galvanic corrosion by completely immersing a joint of dissimilar metals into a polymer, thus avoiding that the metals come in contact with an electrolyte film that could form in presence of humidity.

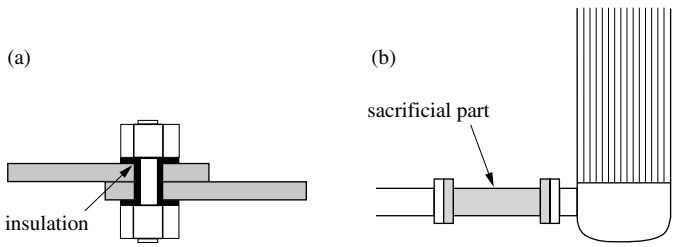


Figure 12.4 Corrosion protection by avoiding galvanic coupling: (a) electrical insulating of a screwed bolt; and (b) use of an easily replaceable sacrificial part.

In large installations, it is not always possible to electrically insulate different materials. The use of sacrificial parts, cheap and easily replaced, allows the problem to be solved in such situations. Figure 12.4(b) shows a sacrificial part made of steel, placed between a brass pipe and a steel heat exchanger.

12.1.3 Selection of materials

When specifying materials for a given project, the corrosion resistance is just one criterion among many others, but it is of uttermost importance for the life time and safety of installations, machines and devices. In this section we focus on the corrosion resistance only and, more specifically, on metallic materials. Engineers in charge of materials selection should keep in mind, however, that for many applications polymers and ceramics present a valuable alternative to metals.

Sources of information

Many sources of information are available that can help engineers to select materials with regard to corrosion resistance:

- internet
- norms and standards;
- specifications of suppliers;
- scientific and technical literature, data bases;
- previous experience with similar equipment;
- laboratory testing.

© 2007, first edition, EPFL Press

At present, the search methods for information on corrosion are rapidly evolving thanks to the development of the internet. A few mouse clicks using well-chosen keywords can yield a wealth of information. While searching the internet is a good starting point, the quality of the information varies enormously and must be carefully checked.

Standards and recommended practices covering different aspects of corrosion have been developed by a number of professional organizations as ASTM, NACE International, ISO and many national bodies. Useful information on the corrosion resistance of materials for a given application can often be obtained from suppliers and from the scientific and technical literature. A number of tables and databases published in the open literature [1- 4] present the intrinsic corrosion resistance of materials in different environments. We must keep in mind, however, that such data are not sufficient, because they do not include electrochemical interactions that can lead to localized corrosion or effects due to stress or wear. Engineers confronted with materials selection therefore need to have a basic understanding of corrosion mechanisms. Experience gained with similar equipment or installations is also a useful source of information to avoid corrosion problems. As a last resort, one may turn to laboratory testing.

Corrosion resistance of metals

Table 12.5 lists in a simplified manner the intrinsic corrosion resistance of pure metals in different environments. It brings out the fact that depending on the environment, one or another of these metals resists better to corrosion. This confirms that corrosion resistance is a property of the metal-environment system and not just of the metal itself.

Table 12.5 Corrosion resistance of some metals.

	Fe	Cr	Ni	Cu	Al	Zn	Ti	Ta
Humid air	–	+	+	+	+	m	+	+
Cold aerated water	m	+	+	+	–	+	+	+
Seawater	–	+	+	m	m	–	+	+
Non-oxidizing acid	–	m	m	+	–	–	+	+
Oxidizing acid	–	+	m	–	–	–	+	+
Acid + Cl^-	–	+	m	–	–	–	m	+
Base + O_2	+	+	+	+	–	–	+	m
High-temperature oxidation	m	+	+	–	–	–	m	–

+ = good resistance
m = average resistance
– = poor resistance

Corrosion resistance of alloys

The corrosion resistance of alloys often resembles that of their principal component. Thus, the corrosion resistance of carbon steel in a given environment is

© 2007, first edition, EPFL Press

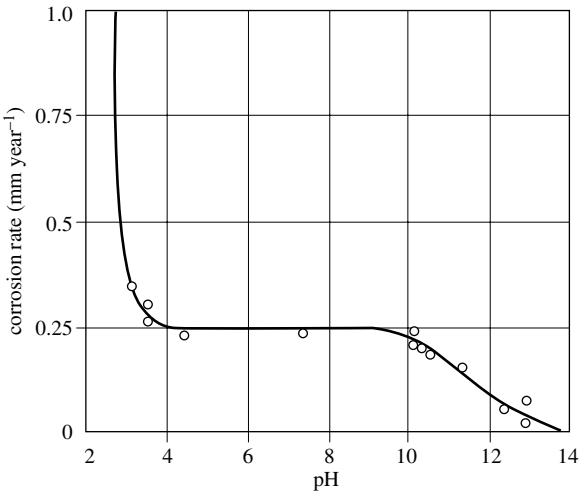


Figure 12.6 Variation of the corrosion rate of steel in water as a function of pH [5].

similar to that of iron, the resistance of aluminum alloys to that of aluminum, etc. In the following, we shall briefly consider the corrosion behavior of :

- carbon steels;
- stainless steels;
- nickel alloys;
- copper alloys;
- aluminum alloys;
- titanium alloys.

Carbon steel

Carbon steels rust when they are in contact with humid air (Chap. 8) and therefore they are usually protected by a coating. In aqueous solutions, their corrosion rate depends on the pH (Figure 12.6). At low pH, proton reduction takes place and the corrosion rate becomes higher as the pH decreases. In neutral solution, oxygen transport controls the rate of corrosion, which therefore does not vary with pH. Finally, in an alkaline solution steel passivates and the corrosion rate decreases to very low values. This explains for example, why steel reinforcements in concrete do not deteriorate as long as the pH stays high (pH $\geq$ 13) but may rust if the pH in the concrete drops to a lower value because of carbonation reactions of cement.

Stainless steel

Chromium in its passive state resists well against corrosion, even in the presence of chloride ions. However, chromium metal is useful only as a coating, because its brittleness renders it unsuitable for bulk applications. Chromium serves mainly as an alloying element in stainless steels. The addition of chromium to steel facilitates the establishment of the passive state in neutral and acidic environments. When the

© 2007, first edition, EPFL Press

chromium content exceeds 12 to 13%, the steel becomes "stainless", meaning that in contact with humid air, a protective passive film forms instead of a rust layer (Chapter 6).

Depending on their structure and their chemical composition, one distinguishes six types of stainless steels (Table 12.7). In addition to the listed alloying elements, stainless steels may contain other elements in small quantity, in particular Cu, Mn, W, Ti, Nb, Ta, P, Si, and N. The different stainless steel families differ not only in their corrosion resistance, but also in their mechanical properties.

Table 12.7 Chemical composition of the different types of stainless steel.

Type	Crystal structure	Cr	Ni	Mo	C	S	N
martensitic	tetrag.	12–18	0–2	> 0.75	0.2	< 0.03	–
ferritic	cc	12–24	0–2	0–0.75	0.1	< 0.03	–
austenitic	cfc	18–20	8–14	0–3	< 0.08	< 0.03	–
super-ferrite	cc	26–29	0.3–4	3–4	< 0.02	< 0.01	0.02
super-austenitic	cfc	20	18–25	6–6.5	0.02	< 0.002	0.2
duplex	cfc+cc	22–26	6–7	4	0.02–0.04	0.001–0.03	0.2

The corrosion resistance of stainless steels increases with their chromium content. Unfortunately, stainless steels with a high chromium content not only cost more, but they are also more difficult to work with (machining, welding) than ordinary stainless steels that contain 13 to 18% chromium. Among these, ferritic stainless steel with 17% chromium (AISI 430), austenitic stainless steel with 18% chromium and 8–10% nickel (AISI 304) and austenitic stainless steel containing 16-18% chromium, 10–14% nickel and 2–3% molybdenum (AISI 316) are most widely used.

Adding molybdenum and nitrogen to a stainless steel improves its resistance to pitting corrosion in chloride-containing environments. Sulfur, on the other hand, present in the form of sulfide inclusions, facilitates pit initiation and therefore lowers the resistance to pitting (Chapter 7). Super ferritic, super austenitic, and duplex stainless steels resist particularly well to pitting corrosion because of their high chromium and low sulfur contents. Figure 12.8 highlights this fact: the pitting potential of a commonly used molybdenum-containing stainless steel (type AISI 316) in seawater at different temperatures is consistently lower than that of a duplex stainless steel (25Cr-7Ni-3Mo-0.14N). The data shown also remind us of the fact that the pitting potential of stainless steels generally decreases with temperature, making them more susceptible to corrosion at higher temperature.

A low carbon concentration reduces the risk of intergranular corrosion of stainless steels that results from the precipitation of chromium carbides at the grain boundaries (Chapter 7). The presence of Ti, Nb or Ta in small amounts (stabilized stainless steels) also works this way, because these elements form particularly stable carbides.

Austenitic stainless steels are sensitive to stress corrosion cracking in chloride-containing environments (Chapter 11). On the other hand, they are generally more resistant to hydrogen embrittlement than ferritic stainless steels.

© 2007, first edition, EPFL Press

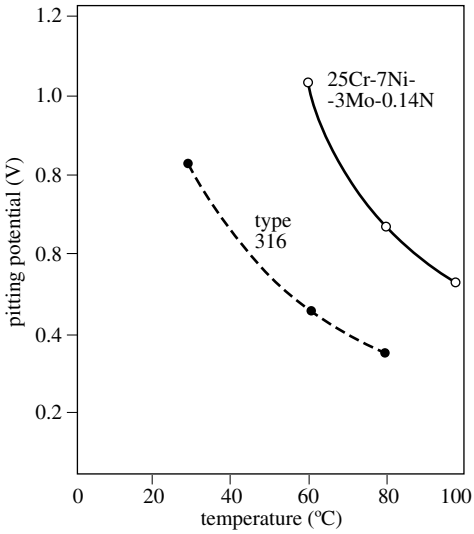


Figure 12.8 Pitting potentials of an austenitic (AISI 316) and a superaustentic (25Cr–7Ni–3Mo–0.14N) stainless steel as a function of temperature in deaerated seawater [6].

Nickel alloys

Nickel exhibits good corrosion resistance in neutral and alkaline solutions and in the atmosphere. However, it corrodes in strongly oxidizing acids. Chloride-containing environments may cause pitting corrosion, especially in commercial grade nickel that contains sulfide inclusions. Pure nickel is quite soft and is used mostly as a coating.

Numerous nickel alloys, in particular those based on copper, chromium, or iron, have both good mechanical properties and high corrosion resistance (Table 12.9). Copper, in particular, improves the resistance against oxidizing acids, while molybdenum improves the resistance towards non-oxidizing acids and chloride-containing environments. Highly corrosion resistant nickel alloys, especially those

Table 12.9 Nickel alloys.

Type	Designation	Ni	Cr	Fe	Mo	W	Cu	Others
Ni	200	99.2		< 0.4				
Ni-Cu	400	ca. 67		1.25			31.5	1 Si
Ni-Mo	B		< 1	5	28			
Ni-Cr-Fe	600	ca. 75	16	8			< 0.5	1.35 Al
	800	32.5	21	44			< 0.75	1 Si
Ni-Cr-Fe-Mo	825	42	21.5	29	3	< 1	2	1 Ti
	G	43	22	19.5	6.5	< 1	2	2 Nb
Ni-Cr-Mo-W	C-276	57	15.5	5.5	16	4		
	C-22	56	22	< 3	13	3		

© 2007, first edition, EPFL Press

suitable for high temperature service, contain substantial amounts of chromium (Chapter 9).

The different nickel alloys are designated by numbers or letters, often preceded by the proper name of the supplier, such as Inconel 600 or Hastelloy C (Table 12.9). Because of their relatively high price, the use of these materials is limited to those applications that require a better corrosion resistance than stainless steels can provide. The corrosion behavior of cobalt alloys is similar to that of the corresponding nickel alloys. Because of their good corrosion resistance some cobalt-chromium alloys are used in biomedical implants.

Copper Alloys

Copper exhibits a good corrosion resistance in air, as well as in hot and cold water, provided the flow velocity does not exceed a certain value (Chapter 10). The standard potential of copper is more noble than that of hydrogen. Therefore, in the absence of oxygen, copper resists corrosion even in acid environments. The potential-pH diagram of Figure 12.10 illustrates this behavior.

Strong oxidizing agents that have a higher standard potential than copper readily attack the metal. As a consequence copper corrodes in HNO_3 and in aerated sulfuric acid. It is rapidly attacked by solutions of $FeCl_3$, a fact that is used in chemical machining of copper. In the presence of H_2S, sparingly soluble copper sulfides precipitate on copper surfaces.

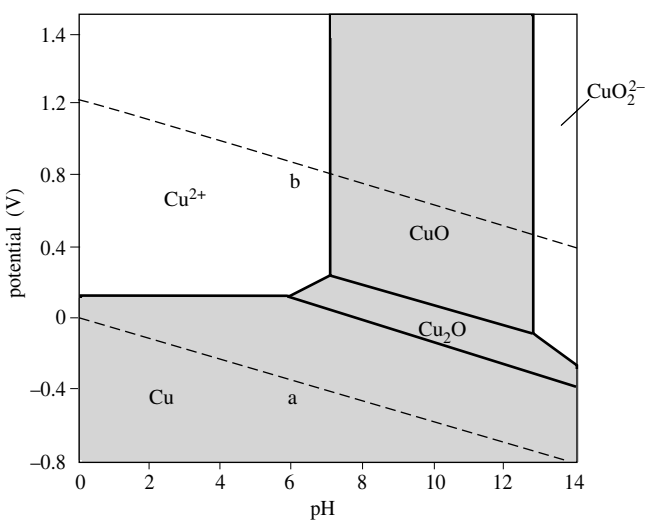


Figure 12.10 The potential-pH diagram for copper. The concentration of dissolved ionic species is 10^{-6} mol l^{-1}. The lines a and b show the stability domain of water.

© 2007, first edition, EPFL Press

Table 12.11 lists three families of copper alloys that find numerous technological applications:

- brass;
- bronze;
- copper-nickel alloys.

The brasses (Cu-Zn alloys) have a corrosion resistance similar to that of copper. Under certain conditions, selective corrosion of zinc may lead to dealloying however (Chapter 7). Adding small amounts of Sn, As, Sb or P to brass permits to reduce this kind of attack. In amine-containing environments, brass is sensitive to stress corrosion (Chapter 11).

Bronzes are copper-tin alloys which upon prolonged contact with the atmosphere form a dark patina that is much appreciated in the art world. In presence of certain pollutants such as chloride the dark patinas eventually turn to green. Aluminum-containing bronze forms surface films containing Al_2O_3 which improves the resistance to erosion corrosion compared to copper or brass.

The copper nickel alloys (Cu-Ni) generally resist well in neutral chloride environments. They are much less sensitive to stress corrosion cracking than brass.

Table 12.11 Copper alloys.

Family	Composition	Examples	Name
Brass	Cu–Zn	Cu–30Zn	Yellow brass
		Cu–30Zn–1Sn	Admiralty brass
		Cu–40Zn–0.75Sn	Naval brass
Bronze	Cu–Sn	C–Sn–P	Phosphorus bronze
		Cu–Al–Ni–Fe–Si–Sn	Aluminum bronze
		Cu–Si–Sn	Silicon bronze
Copper-nickel	Cu–Ni	Cu–30Ni–2Fe	Monel
		65Cu–18Ni–17Zn	Copper-nickel-zinc alloy

Aluminum alloys

Aluminum exhibits good resistance to atmospheric corrosion and to corrosion in neutral solution. The reason is that in this pH range it is protected by a thermodynamically stable oxide film (Figure 2.18). In acidic or alkaline environments, aluminum oxide dissolves and the metal corrodes rapidly with simultaneous formation of hydrogen. In presence of chloride ions, pitting corrosion may occur.

Aluminum alloys show a similar general corrosion behavior as the pure metal, although there may be slight differences related to the alloy composition. The principal alloying elements for aluminum are Fe, Si, Cu, Mg, Mn and Zn (Table 7.41). Among these elements, iron and copper tend to reduce the corrosion resistance, because intermetallic phases can precipitate at the grain boundaries during thermal treatment (Table 7.44), and promote sensitivity to intergranular corrosion and stress corrosion cracking.

© 2007, first edition, EPFL Press

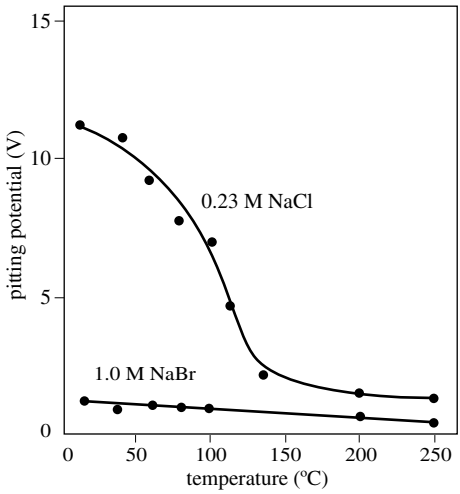


Figure 12.12 Pitting potential of titanium in 0.53 M NaCl and in 1 M NaBr as a function of the temperature [7].

Titanium alloys

Titanium forms extremely stable passive oxide films and therefore exhibits excellent corrosion resistance in oxidizing acids or chloride environments. To illustrate this fact Figure 12.12 shows measured values of the pitting potential in NaCl and NaBr solution at different temperatures. In 0.53M NaCl, the value of E_b at ambient temperature exceeds 10V, which excludes any possibility of spontaneous pitting. At higher temperatures, the pitting potential decreases, but always remains higher than the potential of the oxygen electrode. In reducing environments, titanium is more susceptible to pitting corrosion, because the formation of the passive film requires the presence of an oxidizing agent. To facilitate passivation in non-oxidizing acidic environments one can add to the titanium a small amount of Pd (0.15%) which catalyses proton reduction. The main alloying elements of titanium, however, are aluminum, vanadium and niobium.

A popular titanium alloy, originally developed for aerospace applications, has the composition Ti-6Al-4V. It contains two phases, the α-phase which has a (hc) structure and the β-phase having a (cc) structure. Compared to pure titanium, Ti-6Al-4V exhibits much better mechanical properties, whereas its corrosion resistance in chloride media is almost as good as that of pure titanium. Titanium and its alloys are the materials of choice for biomedical implants that require mechanical strength and good corrosion resistance in chloride media. In addition, titanium and its alloys exhibit good compatibility with biological tissue.

© 2007, first edition, EPFL Press

12.2 COATINGS

12.2.1 Metal coatings

Protection mechanisms

Metallic coatings such zinc and its alloys are commonly used to protect steel against atmospheric corrosion. Copper and brass are often coated with metals for decorative purposes as well as for improved corrosion resistance. Examples include chromium coated water faucets or gold coated electric contacts in printed circuit boards. Metallic coatings also serve for wear protection in tribological applications.

From the point of view of corrosion protection, we distinguish two types of metallic coatings:

- those that are more noble than the substrate;
- those that are less noble than the substrate.

Chromium, nickel and copper used as coatings on steel are more noble than the substrate. In contrast, zinc, cadmium and aluminum form coatings that are less noble than steel. In all cases, the coating provides a barrier between the base metal and the corrosive environment and thus reduces corrosion, provided its own corrosion rate is low. The difference between the two types of coatings becomes apparent in the presence of defects, such as pores or scratches that permit the formation of a corrosion cell between the coating and the substrate.

If the coating is more noble than the substrate (Figure 12.13(a)), the base metal which forms the anode in a corrosion cell experiences galvanic corrosion. The cathode area being much larger than the anode area the effect of galvanic coupling greatly accelerates the local corrosion rate. Coatings that are more noble than the substrate must therefore be free of defects if they are to protect the substrate against corrosion.

On the other hand, when the deposit is less noble than the substrate (Fig. 12.13(b)), the situation reverses. If there is a defect, the coating becomes the anode while the exposed substrate is cathodically protected. Because the surface of the coating largely exceeds that of the defect, the acceleration of the corrosion rate of the coating due to

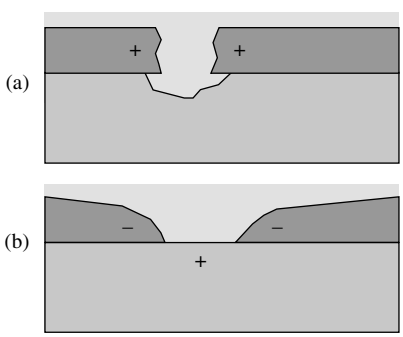


Figure 12.13 Corrosion in the presence of coatings that are (a) more noble and (b) less noble than the substrate.

© 2007, first edition, EPFL Press

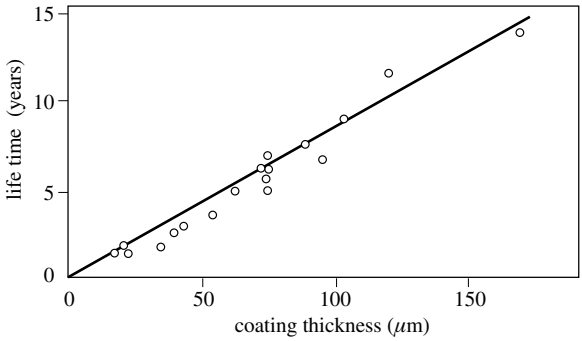


Figure 12.14 Lifetimes of zinc coatings on steel exposed to a polluted atmosphere as a function of their thickness [8].

galvanic coupling remains limited. The useful lifetime of the coating depends on its corrosion rate and is expected to be proportional to the thickness. This has been confirmed for zinc coatings on steel exposed to atmospheric corrosion (Figure 12.14). The lifetimes of zinc coatings fabricated by different methods is plotted as a function of their thickness. In these tests, the lifetime has been arbitrarily defined as the time required for 5% of the total exposed surface to exhibit rusting. Rust appears before the entire coating has been corroded because the cathodic protection of the steel provided by the coating works only over a short distance under typical atmospheric corrosion conditions.

Fabrication of metal coatings

Metal coatings can be fabricated by a variety of processes:

- electroplating;
- electroless plating;
- immersion into liquid metal;
- physical or chemical vapor deposition (PVD, CVD);
- thermal spraying.

To fabricate a coating by ***electroplating***, the substrate is branched as the cathode in an electrolysis cell. The electrolyte contains the depositing metal ions and usually also complexing agents, buffering agents, inert salts and organic additives. Electrodeposition is particularly well suited for coating of large numbers of small pieces such as screws, using continuous barrel plating. High-speed electroplating processes serve for coating of steel sheet in the automobile industry or of lead frames in the electronics industry. Deposition rates as high as a micrometer per second are typical for zinc plating of steel sheet that moves through the electrolyte at speeds that often exceed one meter per second.

A large number of metals can be electrodeposited: Cu, Ni, Cr, Zn, Cd, Sn, Au, Ag, Pd, Pt, as well as many alloys: Sn-Pb, Zn-Ni, Zn-Fe, Au-Cu, Ag-Pd, Ni-Pd, etc. Reactive metals such as aluminum can in principle be electrodeposited from organic

© 2007, first edition, EPFL Press

electrolytes or from molten salts, but such processes have not found any significant application for producing coatings.

Electroless plating is a "reverse" corrosion reaction: metal ions dissolved in the electrolyte are reduced chemically on the substrate surface by a reducing agent that is present in the electrolyte. Electroless plating is widely used to coat steel substrates with Ni-P alloy. It is also used in electronics for the deposition of copper. The reduction of Ni^{2+} ions by hypophosphite ($H_2PO_2^-$) illustrates the principle of the process:

$$Ni^{2+} + H_2PO_2^- + H_2O \rightarrow Ni + H_2PO_3^- + 2\,H^+ \tag{12.1}$$

The actual reaction is more complicated than equation (12.1) because phosphorous and protons are also reduced. As a result, a nickel phosphorous alloy rather than pure nickel is obtained. Under appropriate conditions, the reduction reactions only take place at the metal surface, which acts as a catalyst. Electroless plating thus produces a compact coating of uniform thickness all over the substrate surface. In contrast to electrodeposition, electroless plating can also be used with insulating substrates, such as polymers, the surface of which has been previously activated by a suitable catalyst.

Immersion coating is a classic process for applying zinc onto steel. The object made of steel is briefly immersed into a bath of liquid zinc (boiling point: 420 °C) and a zinc deposit forms whose thickness depends on the immersion time. At the Fe-Zn interface, a interdiffusion zone is created that ensures good adhesion. Other metals with low melting points, in particular aluminum, can be deposited in the same fashion.

Deposition from the gas phase can be performed by a wide variety of physical (PVD) and chemical (CVD) processes, with or without assistance of a plasma. These processes are particularly well suited for the fabrication of thin coatings of refractory materials such as carbides, nitrides, borides or diamond like carbon (DLC). The resulting coatings are useful for tribological and other functional applications, but they generally offer only a limited corrosion resistance.

Physical vapor deposition (PVD) includes simple evaporation as well as sputter deposition processes. The latter are most common. Figure 12.15 schematically

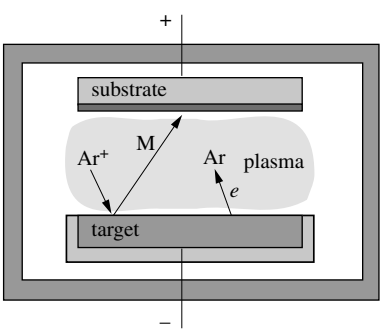


Figure 12.15 DC-current two-electrode sputter deposition system.

© 2007, first edition, EPFL Press

represents a two-electrode, direct-current sputter deposition system, used for the fabrication of metallic coatings. In a vacuum chamber that contains argon or some other gas at low pressure, a voltage on the order of 2000 V is applied between the target (as the negative pole) and the substrate (positive pole). Ionization of the gas results in the formation of a plasma. The electric field accelerates the Ar^+ ions of the plasma in the direction of the target, and the collisions between the impinging argon ions and the target atoms lead to ejection of the latter. They deposit on the substrate (as well as on the walls of the vacuum chamber) where they form a compact adherent coating.

The use of high-frequency alternating current instead of direct current (*rf* sputtering) permits the deposition of non-conducting materials. The difference in mobility between the electrons and the plasma ions creates the electric field required for the deposition. The deposition of compounds such as oxides, carbides, etc., by sputter deposition can be carried out in two fashions, either by using a target made of the compound that is to be deposited or by using a metal target in the presence of a reactive gas such as oxygen, methane, etc. In the latter case, the desired compound forms during the deposition by a chemical reaction between the metal ions and the gas (*reactive sputtering*).

Chemical vapor deposition (CVD) employs a gas mixture that contains a volatile chemical compound of the metal to be deposited and a reducing agent, such as hydrogen or methane. When the gas mixture is introduced into the reaction chamber, the volatile metal compound is catalytically reduced at the surface of the substrate and a thin deposit forms. For example, titanium carbide can be deposited at temperatures between 800 and 1100 °C according to the reaction:

$$TiCl_4 + CH_4 \rightarrow TiC + 4\,HCl \tag{12.2}$$

CVD processes are well suited for the deposition of refractory coatings such as titanium carbides or nitrides, as well as for the fabrication of amorphous or microcrystalline carbon coatings. Conventional CVD processes normally require high temperatures, which limits the number of substrates that can be treated. To circumvent this problem a number of processes have been developed that use a plasma to enhance the reaction rate (plasma enhanced CVD). These processes combine chemical reactions typical for CVD with plasma deposition typical for PVD and therefore the two types of processes increasingly overlap.

Thermal spraying involves spraying of the coating material in liquid form onto a substrate. The droplets rapidly solidify when reaching the surface. The method is particularly well adapted for the fabrication of relatively thick coatings, but unless the process is carefully optimized sprayed coatings risk to be porous. For thermal spraying of low melting metals, such as zinc, the metal to be sprayed is usually introduced into the spray gun as a continuous wire. An oxygen flame provides the energy required to melt the metal and project it onto the substrate to be coated.

Plasma spraying uses a similar approach for producing coatings of refractory materials such as oxides, nitrides or carbides. Because these materials melt at much

© 2007, first edition, EPFL Press

higher temperatures a plasma is formed in the spray gun. The coating material is introduced into the gun as a powder. At the extreme temperatures prevailing in the plasma, the powder melts and partly ionizes. The coating material is accelerated by an electric field established between two electrodes in the plasma gun and is projected onto the substrate where it solidifies. Plasma spraying is widely used for the fabrication of coatings that protect against high-temperature corrosion (Chapter 10).

12.2.2 Inorganic non-metallic coatings

There are two types of inorganic non-metallic coatings: conversion coatings and contact coatings. ***Conversion coatings*** are obtained by a controlled corrosion reaction of the metal to be coated in a suitably chosen solution. Conversion coatings therefore always contain cations from the substrate itself. ***Contact coatings***, on the other hand, are prepared by a deposition process, such as PVD or electroplating, that do not involve a chemical reaction of the substrate. The chemical composition of such coatings is therefore independent of the substrate.

Conversion coatings

Conversion coatings are usually grouped according to their fabrication processes into three types:

- coatings obtained by anodizing;
- coatings obtained by phosphatizing;
- coatings obtained by chromatizing.

The reinforcement of the natural oxide film on valve metals (Al, Ti, Ta, Zr) by anodic oxidation is known as ***anodizing***. Anodizing of aluminum is by far the most

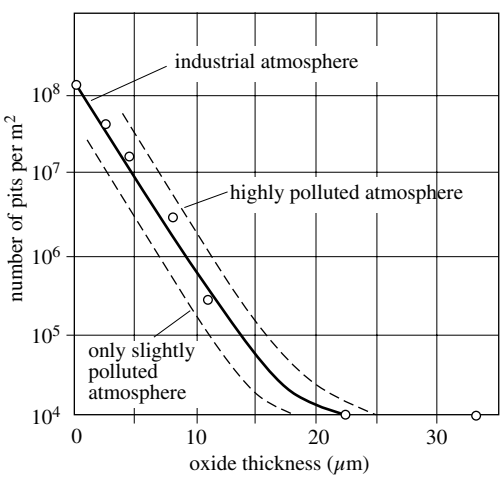


Figure 12.16 Number of pits formed on anodized aluminium as a function of the oxide thickness upon exposition to an industrially polluted atmosphere for a period of 8 years [9].

© 2007, first edition, EPFL Press

important application of this process. Aluminum is often anodized for decorative purposes (coloring), to harden the surface or to increases its resistance towards atmospheric corrosion, in particular pitting. This is illustrated by Figure 12.16 which shows the number of pits formed on anodized aluminum after an exposure to a polluted atmosphere for 8.5 years as a function of the thickness of the anodized film. The resistance to pitting corrosion clearly improves with increasing oxide film thickness.

In order to anodize aluminum, the metal is immersed into an anodizing solution, typically sulfuric acid, and a voltage of 15 to 25 V is applied for a period of about 30 to 60 minutes. The natural oxide film on aluminum grows under the effect of the applied voltage and eventually a thick film is formed, whose structure is schematically shown in Figure 12.17. The metal surface is in contact with a compact inner oxide layer, called the barrier layer, whose thickness varies between 20 and 100 nm depending on applied voltage. The barrier layer grows by high-field ionic conduction as discussed in Chap. 6, the growth constant having a value of 1 to 1.2 nm/V.

The growth mechanism of the outer porous layer involves a dissolution reaction at the bottom of the pores followed by precipitation of oxide near the pore opening. As a consequence, its growth does not require a high electric field and it can reach a much greater thickness than the barrier layer, up to many micrometers. The ionic current passes across the outer film by diffusion and migration of dissolved ions in the electrolyte that fills the pores. The effective ionic conductivity of the outer layer therefore largely exceeds that of the barrier layer.

To seal the pores in the anodic oxide, a second process step is added in which the anodized metal is immersed in boiling water for a few minutes. A hydration reaction occurs, which transforms the alumina of the outer layer into boehmite, a monohydrated aluminum oxide. The resulting increase in volume shuts the pores. The oxide films formed by anodizing can be colored by adding pigments to the anodizing solution. These adsorb onto the surface of the pores in the growing oxide. Pore sealing then traps them inside the pores, yielding a stable colored oxide film.

To reinforce the natural oxide film on other passive metals, in particular stainless steels, a ***chemical passivation treatment*** is sometimes applied. The passivation

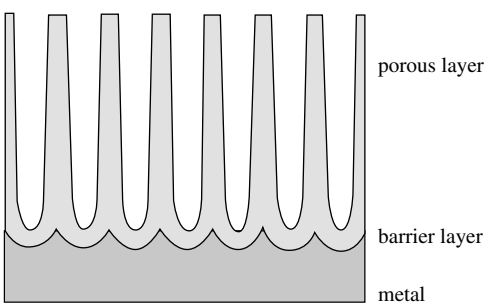


Figure 12.17 Structure of the oxide layer formed by the anodization of aluminium.

© 2007, first edition, EPFL Press

treatment of stainless steel typically involves a short immersion of the metal in nitric acid, an oxidizing agent. This reinforces the passive film, and it also dissolves sulfide inclusions present at the surface, which can act as pit initiation sites. Both effects are beneficial for corrosion resistance.

Phosphatizing is applied to carbon steel in order to improve the adhesion of paints. One distinguishes thin phosphate layers (0.2–0.8 g/m^2), essentially composed of iron phosphate, medium thick phosphate layers (1.5–4 g/m^2), which usually contain zinc phosphate, and heavy phosphate layers (7–30 g/m^2) that contain the phosphates of iron, zinc and magnesium. Heavy phosphate layers provide corrosion protection even in the absence of paint.

Phosphate layers are fabricated by immersing the object into an acid solution that causes corrosion of the steel. The anodic partial reaction yields sparingly soluble iron phosphates which precipitate on the steel surface, thus forming a protecting layer.

$$3\,Fe + 6\,H^+ \rightarrow 3\,H_2 + 3\,Fe^{2+} \tag{12.3}$$

$$3\,Fe^{2+} + 2\,H_2PO_4^- \rightarrow Fe_3(PO_4)_2 + 4\,H^+ \tag{12.4}$$

The actual reaction stoichiometry is more complicated than expressed by the equations 12.3 and 12.4. The reason is that usually an oxidizing agent is added to accelerate the corrosion reaction and, in addition, a variety of precipitation reactions can take place, depending on the electrolyte composition and the operating conditions. In the presence of zinc and magnesium salts, for example, iron-zinc or iron-magnesium mixed phosphates form on the steel surface. Adding a strong oxidizing agent to the electrolyte yields ferric phosphates rather than ferrous phosphates or a mixture of both.

Chromatizing is applied to zinc and to aluminum, often after an anodizing pre-treatment, in order to reinforce the corrosion protection provided by the oxide film. Unfortunately, chromates are toxic and therefore chromatizing treatments are no longer acceptable, in spite of the fact that they provide excellent corrosion protection. They are replaced by less polluting surface treatments such as phosphatizing or suitably chosen coatings.

A chromatizing treatment consists of immersing the object to be treated into an electrolyte that contains chromate or dichromate ions. Reaction (12.5) shows a simplified stoichiometry for the reaction of a bivalent metal with chromates taking place under these conditions.

$$2\,CrO_4^{2-} + 3\,M + 4\,H^+ \rightarrow 3\,MO + Cr_2O_3 + 2\,H_2O \tag{12.5}$$

According to this equation the hexavalent chromate ion is reduced to trivalent chromium oxide thereby oxidizing the metal to its oxide. A surface layer forms that contains both Cr_2O_3 and MO. In reality, the corrosion reactions that take place during chromatizing do not follow such a simple stoichiometry because mixed oxides

© 2007, first edition, EPFL Press

MCr_xO_y, hydrated oxides and hydroxides also form. In addition the surface film contains chromates that have not completely reacted and electrolyte species such as phosphates. The composition of chromatizing layers therefore is rather complex and depends on the operating conditions of the process.

The corrosion protection mechanism of chromatizing layers involves a barrier effect of the oxide film and a chemical effect of the trapped chromates: these maintain oxidizing conditions at the surface and thus facilitate passivation of the metal if pores or other defects are present in the film.

Contact coatings

Non-metallic inorganic contact coatings made of materials not related to the substrate include:

- enamel coatings;
- cement coatings;
- ceramic coatings.

Enamels are silicate-based vitreous coatings. They are used to protect objects made of steel or cast iron. Their thickness varies between 0.05 et 1 mm. They contain SiO_2 (30–60 %), B_2O_3 (10–20 %) and Na_2O (15–20 %), as well as other oxides, in particular Al_2O_3, TiO_2, ZrO_2, CaO, BaO. Enamels are applied in the form of a paste, by immersion or with a spray gun. They are then baked at temperatures generally between 750 °C and 900 °C.

Enamels are corrosion resistant in neutral and acidic environments, even at high temperatures. On the other hand, they degrade more easily in alkaline media. Their behavior varies with composition. A high SiO_2 content, or the presence of TiO_2, increases the corrosion resistance in acids, whereas reducing the concentration of Na_2O, and adding of ZrO_2, improves the resistance in alkaline environments. Enamel coatings protect the substrate by forming a barrier between the metal and the corrosive environment. The protection is therefore only effective as long as the coating is free of defects.

Cement coatings and linings are used in civil engineering to protect structures and pipes made of steel. The coatings are often applied by use of a spray gun. They harden by hydration of the cement.

Refractory ceramic coatings comprise oxides, carbides, nitrides, borides, amorphous carbon and diamond. Thin coatings (typically 3–30 μm), are used for decorative purposes, for improved wear resistance and to a lesser extent for corrosion protection. They are usually applied by PVD techniques as described in the previous section for metallic coatings. Thicker coatings are most often fabricated by plasma spraying. They may reach thicknesses of several hundred micrometer. This type of coating is mostly used to improve the wear and erosion resistance of the substrate material and to provide protection against high temperature corrosion.

© 2007, first edition, EPFL Press

12.2.3 Organic coatings

Classification of organic coatings

Organic coatings used for corrosion protection can be grouped into three categories:

- bituminous coatings;
- polymer coatings;
- paints and varnishes.

Bituminous coatings are used to protect buried steel or cast iron structures against corrosion by providing a barrier between the metal and the environment. Bitumen is composed of a variety of high boiling point hydrocarbons found in the residues of petroleum distillation. The thickness of bituminous coatings can reach several mm. Their plasticity considerably reduces the risk of mechanical damage of the coating that would locally expose the metal.

Polymer coatings for metal protection against corrosion come in a great variety. They include:

- thermoplastic coatings (polyethylene, polypropylene, PVC, PTFE etc);
- rubber coatings (natural or synthetic);
- thermosetting coatings (polyurethanes, epoxies, polyesters etc.).

The thickness of polymer coatings is typically on the order of 100 to 500 μm but varies greatly depending the application. Thick polymeric coatings are applied by a number of techniques: as powder, with spray guns, by lamination, by immersion, etc.

Thermoplastic coatings are often fabricated from powders. To coat a metallic object it is either dipped into a fluidized bed from which the powder deposits or one uses a spray gun to project the powder onto the metal surface. In electrostatic powder coating the polymer powder is projected in the presence of an electric field. This permits coating of recesses and hidden faces that are not in line of sight with the spray gun. After application of the powder, the object is annealed, and the powder transforms into a compact polymer coating.

Thermosetting resins such as epoxies and polyurethanes used as coatings are often produced from liquid precursors. With a spray gun, the two components of the resin are projected onto the metallic surface where they react with each other to form the coating.

The effectiveness of corrosion protection provided by a polymer coating depends on two factors: its intrinsic chemical stability in a given environment and the absence of defects that permit local exposure of the metal. Table 12.18 indicates the intrinsic chemical resistance of a number of thermoplastic coatings.

Paints and varnishes, although chemically similar to polymeric coatings, differ by the way they are applied, which, most often, involves the use of a carrier: water or an organic solvent. However, there are exceptions to this rule: the application of certain oven-dried paints and varnishes used in the automobile industry and elsewhere does not require a carrier. The separation between polymeric coatings and paints is therefore not sharp. Several layers of paint are usually applied to metal surfaces that

© 2007, first edition, EPFL Press

Table 12.18 Chemical resistance of some thermoplasts used as coating materials.

Coating	Useful temperature range (°C)	Acids	Oxidizing acids	Bases	Solvents
PE			–	+	m
PTFE	– 80–250	+	+	+	+
PVC	– 35–65	+	m	m	–
Polyamide (Nylon)	– 50–100	–	–	m	+
Chlorinated polyether (Penton)	– 60–120	+	m	+	+

+ = good
m = medium
– = poor

are to be protected against corrosion. Primary paints form the first layer in contact with the metal and serve for corrosion protection. One or more finishing paints are then applied on top of it.

Paints are normally opaque liquids due to the fact that the polymer forms an emulsion in the solvent. Varnishes, on the other hand, are monophase liquids that may be colored or not. Once applied onto the substrate, they often have a shiny appearance. They are used mostly as decorative layers. In practice, one does not always make a clear distinction between paints and varnishes, and the two terms are often employed interchangeably.

Main components of paints

Four types of components are usually found in paints used for corrosion protection:

- the binder;
- the solvent;
- the pigment,
- the additives and fillers.

The ***binder*** is the most important component of any paint; it guarantees the cohesiveness of the film and its adhesion on the substrate. A wide variety of synthetic polymers as well as derivatives of fatty acids are employed for this purpose, such as PVC (vinyl paints), PMMA (acrylic paints), epoxies, polyurethanes and mixtures of polyesters with fatty acids (alkyd paints).

Paints are usually named after the binder they contain: alkyd paints, vinyl paints, epoxy paints, etc. To obtain specific properties, different polymers are sometimes mixed to form a binder. Table 12.19 shows another classification scheme for paints, based on the modes of drying and reticulation [10]. Certain liquid paints already contain the final polymer chains. During drying (physical drying), van der Waals bonds are created between them as the solvent gradually evaporates. Other paints contain binders of low molecular weight. During drying (chemical drying), these

© 2007, first edition, EPFL Press

molecules react chemically with the air. Another possibility is that they react between themselves in a polymerization reaction (drying by polymerization). In some cases polymerization requires a high temperature and is carried out in an oven.

Table 12.19 Classification of paint according to drying mode

Drying mode	Mechanism of film formation	Examples
Physical drying	Solvent evaporation (organic)	Cellulose paints Vinyl paints Acrylic paints Rubber
	Evaporation of emulsifying agent (water)	Vinyl copolymers Acryl copolymers
Chemical drying	Oxidation polymerization	Alkyds Other binders containing fatty acids
	Polycondensation	Polyesters Aminoplasts Phenyl resins
	Polyaddition	Epoxies Isocyanates
	Radical polymerization	Unsaturated polymers

The presence of a ***solvent***, facilitates the application of a paint in the liquid state. Organic solvents, such as "white spirit", dissolve low-molecular-weight binders. Linseed oil is a natural organic solvent, often used in alkyd paints. Modern paint formulations increasingly use water as a solvent, thus avoiding pollution of the environment and health risks due to solvent evaporation. Water does not dissolve the binder. Rather, the polymer forms a colloidal emulsion, the colloid particles having a diameter of about 0.1 to 0.5 μm. Once the paint has been applied, the water evaporates, leaving behind the binder, which coagulates to a compact film.

Pigments are added to paints to give them a desired color, but a number of pigments serve primarily as corrosion inhibitors. Pigments used for corrosion protection are mostly inorganic compounds with some oxidizing power. Among these, the lead oxide, Pb_3O_4 (minium), and zinc chromate, $ZnCrO_4$, are best known but there are many others: $PbCrO_4$, $CaPbO_4$, $Zn_3(PO_4)$, $SrCrO_4$, $CaMoO_4$, $SrMoO_4$, $ZnMoO_4$, etc.

Minium, Pb_3O_4, of orange color, is widely used in alkyd paints that serve as primary paints on steel for corrosion protection. In a similar way as chromates and molybdates, minium promotes the passivation of the steel surface, although the reaction mechanisms involved are not all known in detail. Unfortunately, pigments containing lead or chromate are toxic and they will have to be replaced with other compounds that are less harmful to the environment. Molybdates or organic molecules such as the tannins can fulfill this task, but their protective power is still inferior. For certain applications ***zinc rich paints***, provide an interesting alternative. They contain

© 2007, first edition, EPFL Press

metallic zinc powder in an organic matrix and can protect steel in a similar way as a metallic zinc coating.

Additives and ***fillers*** fulfill a variety of functions. Some additives modify specific paint properties such as the viscosity, the resistance to oxidation or the surface tension. Fillers are low-cost substances, such as titanium oxide or calcium carbonate, that increase the mass and the volume of paints. They also improve certain properties of the film, such as its abrasion resistance.

Coating systems

To protect metallic structures against corrosion and give them an aspect that appeals to the eye, more than one layer of paint is usually required. In the simplest case, a primary paint containing a corrosion inhibiting pigment is applied for corrosion protection followed by one or more layers of secondary paint to provide a desired color. In applications that require a particularly good corrosion resistance, several layers of different types of coatings are applied. Car body panels provide a good example [11]. Steel may first be coated with zinc or zinc alloy, followed by a phosphate layer to provide good adhesion, followed by a primary coat for mechanical and corrosion resistance, followed by one or two topcoats that provide the desired color and surface finish. In a modern coating system each component has its specific function and its properties must be optimized accordingly.

12.2.4 Protection and degradation mechanisms of paints

Paints protect metallic substrates by different mechanisms, the relative importance of which depends on the chemical nature and the thickness of the coating and on the presence or not of defects and surface contaminants:

- barrier effect;
- inhibition of corrosion reactions;
- prevention of corrosion cells;
- galvanic effect

Barrier effect

Polymer films form a barrier between the metal substrate and the environment. This protects the metal provided that the barrier is impermeable to corrosive agents and remains free of defects. Paint films rarely fulfill these conditions. Firstly, they may contain defects such as pores and scratches that perforate the barrier. Secondly, even intact paint films are permeable to oxygen and water to an extent that varies with their structure, thickness and chemical composition.

The ***permeability*** of a coating for oxygen or water vapor depends primarily on the molecular structure and chemical nature of the binder, especially the degree of cross-linking, but it is also affected by the presence of pigments, additives and fillers. For good corrosion protection, an organic coating should have the smallest permeability possible. To experimentally determine the permeability, one measures the flux through

© 2007, first edition, EPFL Press

a paint film that has been removed from the substrate. The flux N_i of a species i (water, oxygen) through a film of thickness L is equal to:

$$N_i = -D_i \frac{dc_i}{dy} = D_i \frac{c_{i,1} - c_{i,2}}{L} \tag{12.6}$$

In this equation D_i represents the diffusion coefficient of species i in the polymer and c_i its concentration. The concentration at the outside surface, $c_{i,2}$, is given by the solubility of i in the polymer. The concentration at the inside surface, $c_{i,1}$, is zero if the reaction rate at the metal is mass transport controlled. By defining the permeability for species i,

$$P_i^r = D_i \, c_{i,2} \tag{12.7}$$

we obtain for the absolute value of the flux:

$$|N_i| = \frac{P_i^r}{L} \tag{12.8}$$

According to equation (12.8) the flux is inversely proportional to the thickness of the film. A thicker film is therefore less permeable. Solids, such as TiO_2, that are added to paints as fillers or pigments tend to reduce the flux of oxygen or water, provided they exhibit good adhesion to the polymer matrix. The particles increase the effective length of the diffusion path because the diffusing species must pass around them.

At ambient temperature, the solubility of water in a typical paint film exposed to a relative humidity of 100%, reaches 0.5 to 3% by weight and the diffusion coefficient varies between 10^{-14} and 10^{-11} m^2/s [12]. This corresponds to a water permeability of 10^{-8} to 10^{-12} mol/m s. The solubility of oxygen in paints is on the order of some 70 mol/m^3 at atmospheric pressures, and the diffusion coefficient varies between 10^{-12} and 10^{-10} m^2/s [13]. This results in an oxygen permeability of 10^{-8} to 10^{-10} mol m^{-1} s^{-1}. These values indicate that thin paint films do not prevent water and oxygen from reaching the metal surface. However, they slow down the reaction rate.

In spite of being permeable to oxygen and water, paints can protect metals by the barrier effect provided they are impermeable to anions and gaseous pollutants. Atmospheric corrosion is an electrochemical phenomenon (Chapter 8) and requires the presence of an electrolyte at the metal surface. If no anions arrive there, no electrolyte can be formed with the dissolving metal cations and no corrosion takes place.

Inhibition of corrosion reactions

Certain substances added to paint act as corrosion inhibitors. Weakly soluble oxidizing pigments, such as minium (Pb_3O_4) or zinc chromate ($ZnCrO_4$)promote the passivation of the substrate as long as their constituent ions can enter into contact with the metal surface. The presence of water in the film permits to dissolve small amounts

© 2007, first edition, EPFL Press

of inhibitor. These substances are therefore particularly effective in relatively permeable paints, such as alkyd paints.

Passivation is not the only inhibition mechanism possible in a paint film. ***Tannins***, for example, added to certain primary paints for steel, form complexes with the ferric and ferrous ions. The ferric complexes, weakly soluble, precipitate at the film-metal interface, thus reducing the reactive surface area. In addition, the ferrous complexes reduce the concentration of the dissolved ferrous ions. These are required for the formation of magnetite, one of the intermediate steps of the atmospheric corrosion of steel. By preventing the transformation of lepidocrocite into magnetite during consecutive humidity cycles, the average corrosion rate is decreased [14]. In addition, this might reduce the loss of paint adhesion caused by the periodic volume changes associated with these phase transitions.

Suppression of corrosion cells

The presence of an organic coating influences the functioning of microscopic corrosion cells that are responsible for the atmospheric corrosion of steel. In Figure 12.20, an aeration cell in the presence of a paint film is shown. The ionic current between the anode and the cathode either follows the paint-metal interface or passes through the film. In order to reduce the corrosion rate, this current must be kept small. Good adhesion of the paint to the substrate prevents an electrolyte from forming at the paint-metal interface and thus augments the electrical resistance between anodic and cathodic areas. A low ionic conductivity of the paint matrix reduces the current flow within the film. The ***electrical resistance*** of a paint is therefore a valid quality criterion. Its value should be between 10^{13} and 10^{15} ohm m^{-2}, but in no case lower than 10^{11} ohm m^{-2} [12].

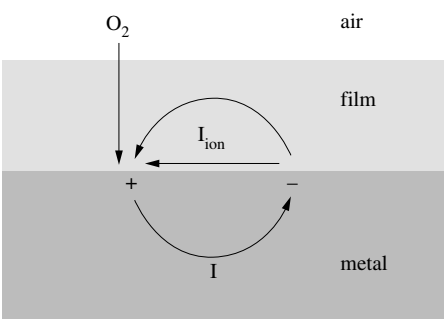


Figure 12.20 The electric and ionic currents in a corrosion cell in the presence of an organic coating. The ionic current passes through the film or along the film-metal interface. The electrons in the metal move in the direction opposite to the current.

Galvanic effect

Zinc rich paints contain up to more than 90% zinc powder, mixed with an organic resin. They are mostly used as primary coats on steel to provide corrosion protection.

© 2007, first edition, EPFL Press

The protection mechanism is similar to that of metallic zinc coatings: zinc is less noble than steel and protects the substrate by forming a galvanic corrosion cell, in which zinc is the anode. This mechanism is particularly active in presence of defects in the coating that provide an electrolyte path to the substrate surface. In addition, the atmospheric corrosion of zinc yields voluminous solid corrosion products (oxides, carbonates, etc.) that are capable of blocking pores or small defects in the coating, thereby reinforcing its barrier effect.

Degradation of paint films

Several mechanisms can contribute to paint degradation and to the loss of its protective properties (Figure 12.21):

- ageing of the resin;
- formation of blisters;
- cathodic delamination.

Ageing

The ageing of paints (Fig. 12.21(a)) is the result of complex physical and chemical changes in the polymer that take place upon prolonged exposure to the environment. Ultraviolet radiation and oxidation reactions are particularly damaging. They render the paint film increasingly brittle and thus reduce its resistance to abrasion and to cracking. The defects thus created facilitate the formation of aeration cells.

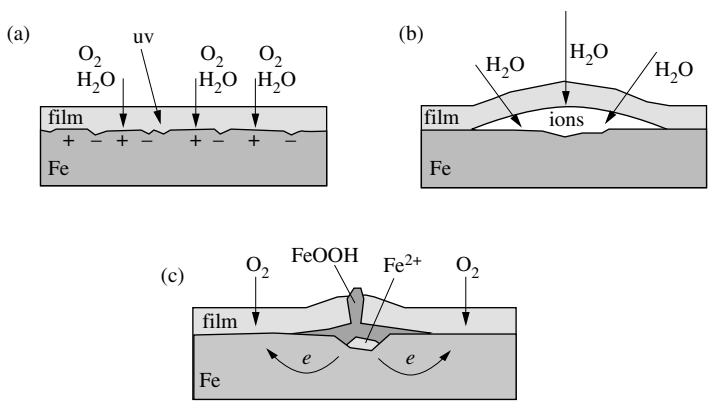


Figure 12.21 Mechanisms of the degradation of paint films: (a) degradation of a polymer due to ultraviolet radiation or chemical attack; (b) the formation of blisters by osmosis; and (c) cathodic delamination due to the formation of corrosion cells.

Blister formation

Blister formation (Fig. 12.21(b)) is an osmotic phenomenon in which the paint film acts as a semi-permeable membrane. Ionic contaminants present at the interface between the metal and the paint film form a concentrated electrolyte with the water

© 2007, first edition, EPFL Press

diffusing through the film. An osmotic pressure develops that attracts more water. The film locally detaches from the substrate and a small blister forms. If the film exhibits high plasticity and has a relatively low adhesion to the substrate, the blisters grow with time and an ever-larger fraction of the paint film detaches from the substrate. On the other hand, in case of a brittle film, the pressure build up at the interface causes cracking of the film and exposure of the substrate. In both cases the paint looses its protective properties.

Cathodic delamination

Cathodic delamination is a loss of adhesion that results from an increase in pH due to a cathodic reaction at the metal-coating interface. Cathodic delamination as well as blister formation are often the result of a faulty application of the coating, in particular insufficient cleaning of the metal before painting. Defects in the film that result from mechanical damage (scratches, impacting particles) are also a starting point for cathodic delamination, because corrosion cells may form between the protected and damaged areas.

Figure 12.21(c) schematically shows the reactions that cause cathodic delamination during the atmospheric corrosion of a coated steel. In presence of a defect, the steel surface becomes exposed to the atmosphere and the metal locally corrodes. Because oxygen can diffuse through the paint film it reacts not only at the defect site, but also at the paint-metal interface nearby, rendering it slightly cathodic with respect to the defect site. A corrosion cell starts to form. As rust increasingly covers the defect area, the access of oxygen to the metal surface diminishes and this accentuates the development of the corrosion cell. The defect area thus becomes increasingly anodic relative to the paint-covered metal. The reduction of oxygen underneath the paint (cathodic partial reaction) leads to an increase of the pH at the metal-film interface and, as a consequence of this, to a loss of adhesion of the paint and its flaking off (cathodic delamination). A pH increase at the paint-metal interface, and hence cathodic delamination, is also observed when a coated metal is subjected to an intense cathodic polarization. The same mechanism accounts for loss of adhesion during blister growth.

The loss of adhesion resulting from the rise in pH during cathodic delamination has been explained by different phenomena: saponification reactions of functional groups of the binder (esters, amides); reactions of intermediates of the oxygen reduction (H_2O_2, OH^-) with the coating; dissolution of the oxide at the metal-paint interface rendered possible by the high pH. In practice, the risk of cathodic delamination can be reduced by improving the chemical resistance of the paint film to alkaline attack or by an appropriate chemical pretreatment of the metal surface such as phosphatizing in order to stabilize its pH.

12.2.5 Experimental study of metal protection by organic coatings

Accelerated exposure tests

Exposure tests, accelerated or not, are still most widely used in industry for evaluating the durability of organic coatings and their effectiveness for corrosion

© 2007, first edition, EPFL Press

protection. These tests are described in detail in numerous norms and standard practices edited by professional organizations. ***Accelerated exposure tests*** are based on the use of climate chambers that permit to simulate wetting and drying cycles in a controlled atmosphere (pollutants), with or without periodic irradiation of the sample with light or with ultraviolet radiation. Salt spray tests are also popular for the evaluation of certain types of coatings. The aggressiveness of the test can be increased by adding oxidizing agents such as a copper salt to the salt mist. Because corrosion of coated metals usually starts at defects such as scratches or pinholes, one can accelerate testing by using coatings that contain a scribe. Corrosion then progresses from the scribe and can be easily located. Using scribed samples also improves the reproducibility of corrosion tests of coated metals.

Electrochemical methods

Electrochemical methods are well adapted for characterizing the corrosion behavior of coated metals in solution. Because of the high resistance of organic coatings, ac methods are generally more suited than dc polarization methods. In electrochemical impedance spectroscopy (EIC) one measures the response of the coated electrode to a small amplitude *ac* perturbation as a function of frequency (Chapter 5). The interpretation of the measured frequency response, in principle, requires a physical model. However, for coated metals useful information is more easily obtained by representing the metal-coating-electrolyte interface by an electrical circuit (equivalent circuit).

A simple equivalent circuit that provides an intuitively reasonable description of the different contributions to the impedance is shown in Figure 12.22. It assumes that the coating, which is non-conducting, contains pores and pinholes filled with electrolyte through which the current flows. The circuit contains three resistances: the electrolyte resistance between the reference electrode and the coating, R_{soln}, the resistance in the pores of the coating, R_{pore}, and the charge transfer resistance at the interface metal-electrolyte at the bottom of the pores, R_t. In parallel to the charge transfer resistance is the double layer capacity at the metal-electrolyte interface in the

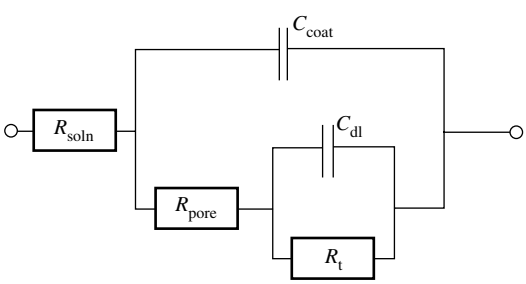


Figure 12.22 Electrical model circuit for the interpretation of the impedance response of a coated metal in contact with an electrolyte.

© 2007, first edition, EPFL Press

pores, C_{dl}. Finally, the capacity C_{coat} represents the capacity of the coating. In the absence of pores, the interface would behave as a capacitance of value C_{coat}.

The variation of the coating capacity with time yields the rate of water uptake of a coating because water changes the dielectric constant and hence the value of C_{coat}. The variation of R_{pore} with time provides an indication of the rate of deterioration of the coating: R_{pore} diminishes progressively as the coating looses adhesion. However, the relation between R_{pore} and delaminated area is not straightforward and therefore only qualitative information is obtained. A serious difficulty with impedance measurements applied to the study of coatings, is reproducibility. To improve the accuracy of results, one can use statistical methods for data interpretation or study coatings that contain well-defined artificial defects.

Adhesion tests

Good ***adhesion*** of the coating on the substrate is crucial for efficient corrosion protection. There are several methods for measuring the adhesion of a coating, but none of them really satisfactory. In the field, one measures the adhesion of coatings with an adhesive tape. A chosen area of the coating is cut into small squares with a knife. An adhesive tape is brought into contact with the cut area and then torn off. Some squares will stick to the tape, others remain on the substrate. The relative proportion of these is a qualitative measure of adhesion.

To measure adhesion in the laboratory, a small metal disk is glued to the coating. After drying, the assembly is subjected to a tensile test and one measures the force needed to separate the coating from the metal. Depending on the adhesive strength of the interface with respect to the cohesive strength of the coating, the fracture surface will lie at the coating-metal interface or within the coating itself. In a similar way, one can apply a torsion force instead of a traction. Both methods permit in principle to quantify adhesion, but they have serious shortcomings that limit the reproducibility of obtained results: the alignment of the set up during the test is critical, and gluing of the coating to the metal disk must not alter its properties.

Physical methods for the study of the coating-metal interface

In recent years a number of physical methods have been developed for the study of corrosion reactions underneath coatings including the scanning Kelvin probe and scanning acoustic microscopy. These methods are not yet used for industrial testing, but they are useful for studies aimed at a better understanding of reaction mechanisms that govern the effectiveness of coatings for corrosion protection and their adhesion to the substrate.

Scanning Kelvin Probe

The Scanning Kelvin Probe (SKP) serves to measure changes in the corrosion potential of film covered metals [15]. The same method is applicable to the study of

© 2007, first edition, EPFL Press

potential changes during wet-dry cycles in atmospheric corrosion (Figure 8.22). The principle is the following.

When two metals that differ in their work function (and hence in the electrochemical potential of their electrons, cf. Chapter 2) are brought into electrical contact, a small amount of charge passes from one to the other, until the electrochemical potentials of the electrons become equal. The excess charge establishes a potential difference $\Delta\Psi$ (contact potential) that is equal to the original difference in work function.

$$F\,\Delta\Psi = \Phi_{W,1} - \Phi_{W,2} \tag{12.9}$$

In this equation $\Phi_{W,1}$ is the work function of metal 1 and $\Phi_{W,2}$ that of metal 2. To measure the potential difference $\Delta\Psi$, one uses a vibrating probe, schematically presented in Figure 12.23. The sample and the Kelvin probe are separated by a dielectric and thus form a capacitor:

$$C = \varepsilon\,\varepsilon_0\,\frac{A}{L} \tag{12.10}$$

Here C is the capacity, ε is the dielectric contact, ε_o the permittivity constant, A the surface area and L is the distance between the capacitor plates. Because the probe is vibrating the distance L varies periodically: $L = L_0 + \Delta L\,\sin(\omega t)$. The variation of L with time is reflected in the capacity C. The variation of the capacity induces a current:

$$I = \Delta\Psi\,\frac{dC}{dt} \tag{12.11}$$

For a sinusoidal variation this yields:

$$I = \Delta\Psi\,\varepsilon\,\varepsilon_o A\,\omega\,\Delta L\,\frac{\cos(\omega t)}{(L_0 + \Delta L\sin(\omega t)^2)} \tag{12.12}$$

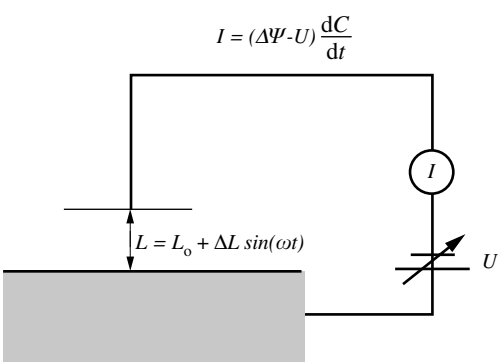


Figure 12.23 Principle of the Kelvin probe; the sinusoidal variation of the distance between the probe and the sample induces a capacitive current. When the applied voltage U equals $\Delta\Psi$ the current is zero.

© 2007, first edition, EPFL Press

which for dL << L reduces to

$$I = \Delta\Psi\, \varepsilon\varepsilon_o A\, \omega \Delta L \frac{\cos(\omega t)}{L^2} \tag{12.13}$$

To determine $\Delta\Psi$ one applies a variable dc voltage of opposite sign while simultaneously measuring the current,

$$I = (\Delta\Psi - U)\frac{\mathrm{d}C}{\mathrm{d}t} \tag{12.14}$$

and one varies U until the current drops to zero when $U = \Delta\Psi$. For measurements carried out in vacuum, the potential difference $\Delta\Psi$ has a clear physical meaning, but its interpretation in presence of an electrolyte is not straightforward**.** Experimental measurements in electrolytes indicate that $\Delta\Psi$ is proportional (but not equal) to the corrosion potential [16]. Using a calibration procedure the Kelvin probe is therefore capable to determine changes in corrosion potential of a coated metal or of a metal exposed to a corrosive atmosphere. By using a sufficiently small probe and scanning it over the sample surface, local potentials can be determined permitting the establishment of potential maps on a surface. Because the scanning SKP measurements are very sensitive to distance (Equation 12.13) the surface should be flat or, alternatively, the SKP measurement corrected for surface topography.

Scanning acoustic microscopy

Scanning acoustic microscopy (SAM) permits the study of disbonding at the coating-metal interface. The principle is the following (Figure 12.24)**:** An acoustic signal in the 0.1 – 1GHz range is focused on the coating/metal interface by means of a sapphire lens immersed in the electrolyte. For ex situ studies one uses a water drop placed on the coating. The acoustic waves are generated by a piezoelectric transducer

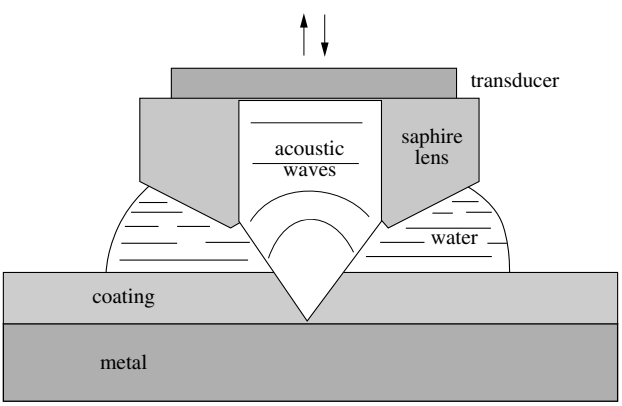


Figure 12.24 Schematic presentation of the acoustic microscope: the reflection of the acoustic waves from the coating-metal interface depends on the interface properties [17].

© 2007, first edition, EPFL Press

which also acts as receiver for the reflected waves. The reflection contrast results from differences in the elastic properties of the phases, or more precisely, in their acoustic impedance. The acoustic impedance is defined as $Z_{ac} = v_{ac}\,\rho$, where v_{ac} is the velocity of the acoustic wave in a given material and ρ is its density. The reflection coefficient of the paint-metal interface, R_f, is given by [17]:

$$R_f = \frac{Z_{ac,2} - Z_{ac,1}}{Z_{ac,2} + Z_{ac,1}} \tag{12.15}$$

where the subscripts 1 and 2 indicate, respectively, the coating and the metal. Water or electrolyte solution present at the interface during disbonding leads to a change in reflectivity. By scanning the lens over the coated surface an image of the advancement of disbonding can therefore be obtained. Because the image is based on acoustic and not on optical waves the method can be applied equally well to transparent and to opaque coatings. In a similar way as the Kelvin probe, SAM requires flat surfaces because the signal intensity is quite sensitive to distance.

Spectroscopic methods

Vibrational spectroscopies, such as FTIR or Raman spectroscopy can be applied to the study of chemical changes taking place in coatings as a result of exposure to a corrosive environment. ***Surface analysis methods*** such as AES, XPS and SIMS (Chapter 3) or glow discharge spectroscopy (GDS) provide information on corrosion induced chemical changes at the coating-metal interface. For this, one can either sputter through the interface (depth profile analysis), or one can analyze the contact surfaces after the coating has been detached from the metal.

12.3 CORROSION INHIBITORS

12.3.1 Classification of corrosion inhibitors

A ***corrosion inhibitor*** is a chemical compound that is added to the environment in order to reduce the rate of metal corrosion. The ***inhibition efficiency***, R_I, indicates to what extent the corrosion rate is slowed down by the presence of the inhibitor:

$$R_I = \frac{v_0 - v}{v_0} \tag{12.16}$$

Here, v_0 and v designate the rates of corrosion in the absence and in the presence of the inhibitor, respectively. The inhibition efficiency depends on a number of parameters, an important one being the inhibitor concentration.

Inhibitors can be classified in different ways (Figure 12.25):

- by their field of application;
- by their effect on the partial electrochemical reactions;
- by their reaction mechanism.

© 2007, first edition, EPFL Press

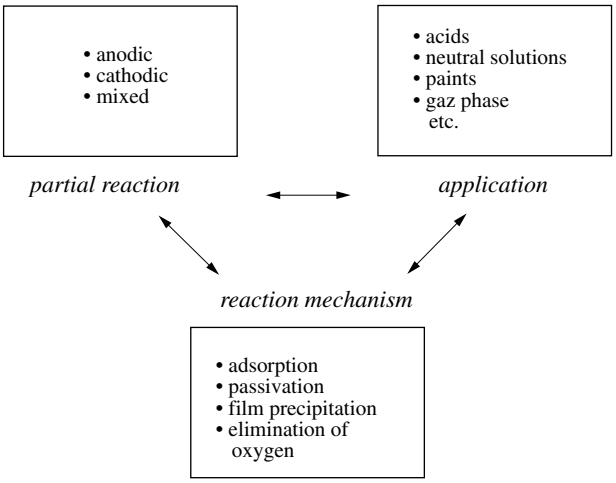


Figure 12.25 Different ways of classifying corrosion inhibitors.

Fields of application

In practice, inhibitors are often defined according to their field of application. In aqueous environments, ***inhibitors for acid environments*** are typically used to minimize metal corrosion during pickling of steel, an operation that removes oxide scales by dissolution in an acid. In the petroleum industry, large quantities of inhibitors for acid environments are used to avoid corrosion of drilling equipment. ***Inhibitors for neutral environments*** are used above all for the protection of cooling-water circuits. Inhibitors not only reduce the rate of uniform corrosion, but they also serve to protect metals from localized corrosion and stress corrosion cracking [18].

Corrosion inhibitors are used also in organic fluids such as lubricating oils and gasoline. These liquids often contain traces of water and ions that can cause corrosion. ***Corrosion inhibitors for paints***, include both inorganic pigments and organic molecules such as for example derivatives of tannic acid (Section 12.3).

Gas-phase inhibitors serve for the temporary protection of packaged objects during transport and storage: machines, precision instruments, electronic components, etc. Gas phase inhibitors typically are organic compounds of low molecular weight with a high vapor pressure. These compounds adsorb from the gas phase onto the metal surface and thus reduce the rate of atmospheric corrosion. Dicyclohexylamine nitrite (Figure 12.26), used for the protection of steel, is a typical example of such an inhibitor. Its vapor pressure is 1.3×10^{-3} Pa at ambient temperature. Another way to protect packaged objects from atmospheric corrosion is based on the use of ***dessicants*** such as silica gel: the presence of these compounds reduces the relative humidity of the air inside the package and thus prevents corrosion.

© 2007, first edition, EPFL Press

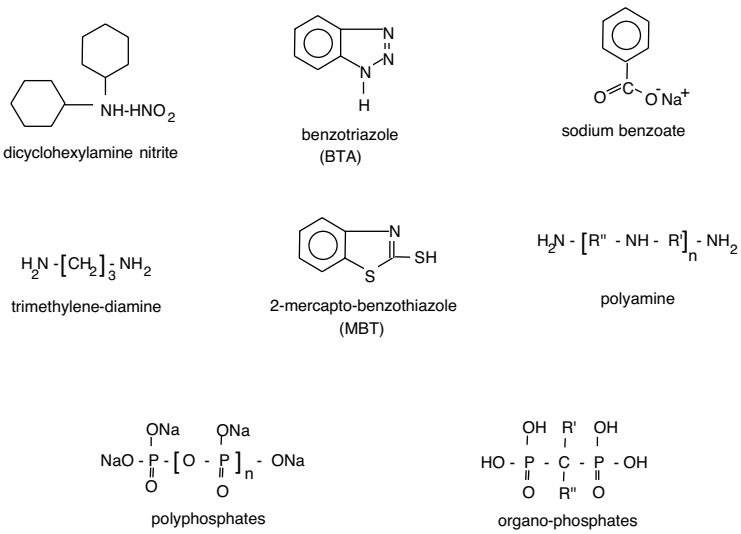


Figure 12.26 Chemical formulas of several corrosion inhibitors.

Effect on partial electrochemical reactions

The presence of a corrosion inhibitor reduces the rate of corrosion by affecting the rate the partial electrochemical reactions involved. Depending on which partial reaction is mostly slowed down one distinguishes three types of inhibitors:

- anodic inhibitors;
- cathodic inhibitors;
- mixed inhibitors.

Figure 12.27 schematically shows how the presence of corrosion inhibitors affects the polarization curve of systems in which the partial reactions are governed by Butler-Volmer kinetics. A ***cathodic inhibitor*** lowers the cathodic partial current

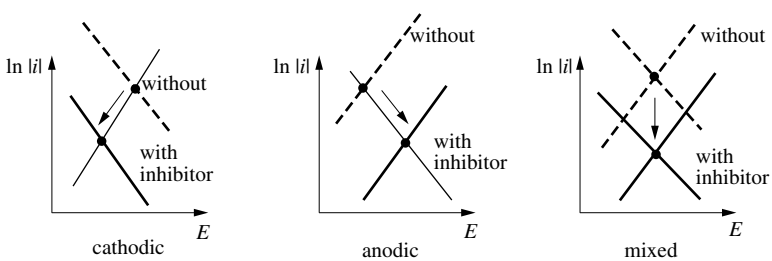


Figure 12.27 Evans diagrams showing the shift of the corrosion potential due to the presence of an anodic, cathodic or mixed inhibitor.

© 2007, first edition, EPFL Press

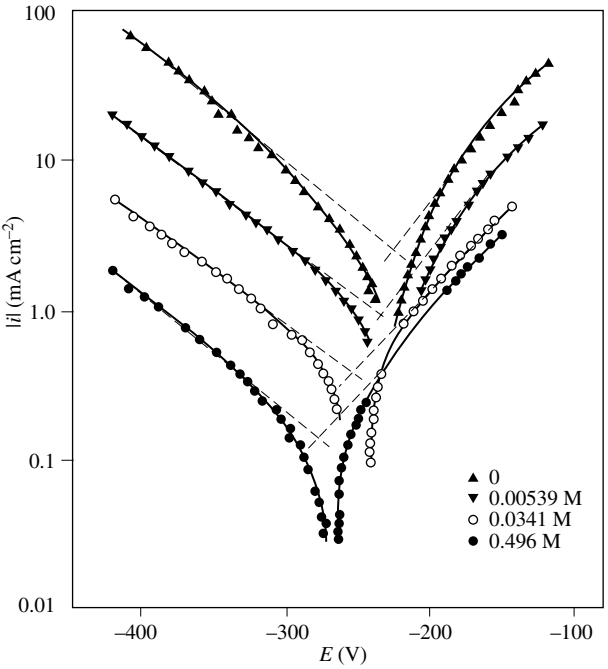


Figure 12.28 Polarization curves for iron in 6 M HCl in the presence of different concentrations of trimethylene-diamine corrosion inhibitor [19].

density and shifts the corrosion potential in the negative direction, whereas an ***anodic inhibitor*** reduces the anodic partial current density and shifts the corrosion potential in the positive direction. A ***mixed inhibitor*** reduces the rate of both partial reactions, but has little effect on the corrosion potential.

The polarization curves for iron, measured in 6M HCl solutions containing different amounts of trimethylene diamine (Figure 12.28), illustrates the described behavior [19]. This compound, whose formula can be found in Figure 12.26, reduces the cathodic partial current density more than the anodic partial current density and its presence lowers the corrosion potential. On the other hand, the slope of the Tafel lines is not affected by the inhibitor, indicating that the mechanism of the electrode reaction remains unchanged. The reduction of the partial currents can be explained by postulating that the inhibitor adsorbs on the metal and thus reduces the surface area available for the corrosion reactions.

The distinction between anodic and cathodic inhibitors is not limited to Butler-Volmer kinetics, but applies to all types of reaction kinetics. For example, a cooling water inhibitor that, by formation of a porous film, reduces access of oxygen to the metal surface is a cathodic inhibitor. On the other hand, an inhibitor that reduces the rate of corrosion of a metal by promoting its passivation, is an anodic inhibitor.

© 2007, first edition, EPFL Press

Reaction mechanisms of inhibitors

Corrosion inhibitors act by different mechanisms:

- adsorption;
- passivation;
- film formation by precipitation;
- elimination of the oxidizing agent.

The ***adsorption*** of an inhibitor onto the metal surface slows the rate of corrosion by blocking part of the surface. The extent of inhibition depends on the equilibrium between the dissolved and adsorbed inhibitor species, expressed by the adsorption isotherm. This mechanism which is particularly important in acids will be discussed in the next section. (Sect. 12.4.2). Certain inhibitors promote the spontaneous ***passivation*** of a metal and thus drastically reduce the corrosion rate. Oxidizing species such as chromates fall in this category. Buffer agents that maintain a high pH at the metal surface also favor the passive state. Other inhibitors lead to the formation of surface films by ***precipitation*** of mineral salts or of weakly soluble organic complexes. These films reduce the ability of oxygen to reach the surface and, in addition, they may impede the anodic dissolution reaction.

Inhibition by elimination of the oxidizing agent is only applicable in closed systems. So-called ***oxygen scavengers*** are used to eliminate the last traces of oxygen in hot water circuits of power plants. For this a small amount of sodium sulfite or hydrazine is added to the water, which has been previously degassed and deionized:

$$SO_3^{2-} + \frac{1}{2}\,O_2 \rightarrow SO_4^{2-} \tag{12.17}$$

$$N_2H_4 + O_2 \rightarrow N_2 + 2H_2O \tag{12.18}$$

By eliminating the oxygen one also eliminates corrosion. We have seen previously that adding a desiccating agent to a closed atmosphere reduces atmospheric corrosion by eliminating humidity. This is another way of using an inhibitor for reducing the aggressiveness the environment.

12.3.2 Inhibitors for acid environments

A large number of organic substances, most often aromatic compounds or macromolecules with linear or branched chains, serve as corrosion inhibitors for acid environments. These inhibitors adsorb onto the surface of the metal and thus slow the rate of corrosion. In some cases film formation can occur. The inhibition efficiency in a given application depends on the molecular structure of the inhibitors and on their concentration.

Molecular structure of adsorption type inhibitors

An organic inhibitor molecule typically includes a hydrophobic non-polar part that contains mostly carbon and hydrogen atoms, and one or several hydrophilic

© 2007, first edition, EPFL Press

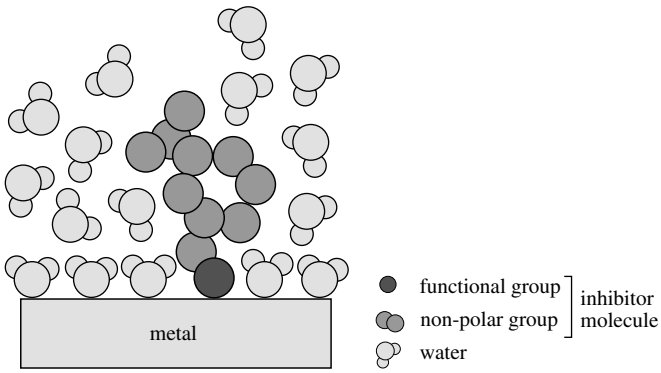


Figure 12.29 Adsorption of an organic inhibitor onto a metal surface in aqueous environment.

functional groups such as -NH_2 (amine), -SH (mercapto), -OH (hydroxyl), -COOH (carboxyl), -PO_3 (phosphate) or their derivatives (Figure 12.26). The polar functional groups permit dissolution of the inhibitor in water and they link the inhibitor molecule to the metal surface. The non-polar part, which takes up a much greater volume, partially blocks the active surface area (Figure 12.29). The described behavior is not limited to acid solutions, but typical for all types of adsorbing inhibitors. For example, trimethylene diamine which is a gas-phase inhibitor and sodium benzoate or organophosphates that are used in neutral solution follow the same principle.

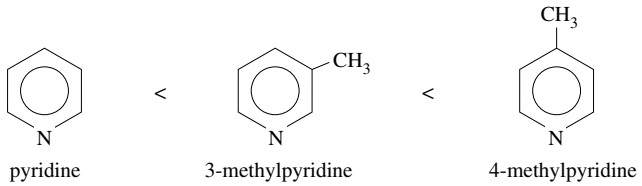
Influence of the electron density

Inhibitors for acid environments generally chemisorb on the surface. The effectiveness of the inhibitor therefore is the better, the greater the tendency of the functional groups to form bonds with the metal by acting as electron donors. For organic molecules that differ only by their functional groups, the inhibition efficiency varies inversely with the electronegativity of the functional atoms. It increases in the following order [19,20]:

$$O < N < S < Se < P$$

The molecular structure of the non-polar part can also influence the ability of the functional atoms to donate electrons. By placing a methyl group, -CH_3, at the 3 or 4 position of a pyridine molecule, the electron density at the nitrogen atom is increased, thus promoting a bond with the metal, the reason being that the methyl group repels electrons (nucleophilic group). The inhibition efficiency of pyridine derivatives thus increases in the following order [19]:

© 2007, first edition, EPFL Press



If, on the other hand, instead of a methyl group the above molecules contain an electrophilic substituent that attracts electrons, such as –Cl, the inhibition efficiency decreases.

In the same way, the effectiveness of inhibition by benzonitrile, C_6H_5-CN varies with the presence of nucleophilic or electrophilic groups. Figure 12.30 shows the corrosion rate of iron in a solution of 0.5 M H_2SO_4 (25 °C), in the presence of 0.01 mole/l of different benzonitrile derivatives. The parameter σ (Hammet constant), plotted on the abscissa, expresses the electron density at the group –CN; a large positive value of σ corresponds to a low electron density. The results reveal a semi-logarithmic correlation between the corrosion rate and the parameter σ. The corrosion rate is lowest in presence of benzonitrile with a the nucleophilic substituent $-CH_3$ giving the highest electron density on the –CN group.

The general picture given here applies only to the simplest kinds of molecules. Often other factors, in particular steric effects, influence the adsorption behavior of organic molecules. In that case, there is no simple correlation between the electron density at the functional group and the inhibition efficiency. The same is true if the inhibitor forms surface films of more than one atomic layer, or if it undergoes an oxidation or reduction reaction at the surface. An even more complicated situation arises if the inhibitor undergoes a chemical transformation and the reaction product adsorbs on the metal, a mechanism referred to as *secondary inhibition*.

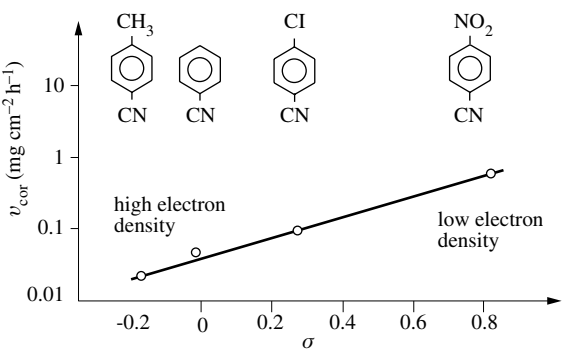


Figure 12.30 Variation of the corrosion rate of iron in 0.5 M H_2SO_4 as a function of the σ parameter, which characterizes the electron density at the –CN functional group of the inhibitors (of the benzonitrile family) shown in the figure [20].

© 2007, first edition, EPFL Press

Influence of the concentration

The adsorption isotherm describes the variation of the surface coverage of an adsorbing species as a function of its concentration in the gas or liquid phase present. If the adsorption of an inhibitor from solution follows the Langmuir isotherm (Chapter 3) its surface coverage θ is given by:

$$\theta = \frac{b_L c_{inh}}{1 + b_L c_{inh}} \tag{12.19}$$

In this equation b_L designates the adsorption coefficient and c_{inh} is the inhibitor concentration in the electrolyte.

A simple model allows us to relate the corrosion current density of a metal to the inhibitor surface coverage. We assume that the corrosion current density in the presence of an inhibitor is the sum of two terms:

$$i_{cor,\theta} = (1-\theta) i_{cor,\theta=0} + \theta\, i_{cor,\theta=1} \tag{12.20}$$

Here (1-θ) is the fraction of non-covered surface and $i_{cor,\theta=0}$ represents the corrosion current density observed in the absence of inhibitor. The term $i_{cor,\theta=1}$ represents the corrosion current density measured on a surface entirely covered by the inhibitor. After rearrangement of (12.20), we obtain:

$$\theta = \frac{i_{cor,\theta=0} - i_{cor,\theta}}{i_{cor,\theta=0} - i_{cor,\theta=1}} \tag{12.21}$$

With equation (12.21) we can determine θ by measuring the rate of corrosion as a function of the inhibitor concentration.

Another way to determine θ is based on the measurement of the ***double layer capacity***, whose value is smaller in the presence of an adsorbed inhibitor. As long as the contributions of the covered and uncovered parts of the surface are simply additive, the double layer capacity of a partially covered electrode C_θ is equal to:

$$C_\theta = (1-\theta) C_{\theta=0} + \theta C_{\theta=1} \tag{12.22}$$

where $C_{\theta=0}$ represents the double layer capacity in the absence of inhibitor, and $C_{\theta=1}$ designates the capacity of an entirely covered surface. After rearrangement, we obtain:

$$\theta = \frac{C_{\theta=0} - C_\theta}{C_{\theta=0} - C_{\theta=1}} \tag{12.23}$$

The results given in Figures 12.31 and 12.32 illustrate the application of these relations to the identification of an inhibition mechanism. Figure 12.31 shows the corrosion current density of iron as a function of benzotriazole concentration added to sulfate solutions of different pH [21]. From these results, the surface coverage of the

© 2007, first edition, EPFL Press

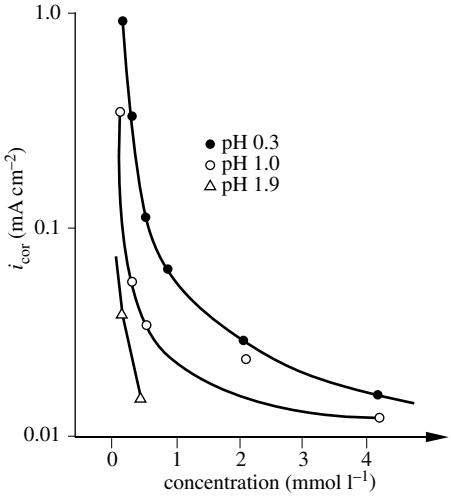


Figure 12.31 Effect of the benzotriazole concentration on the corrosion rate of iron in sulfate solutions of different pH [21].

inhibitor can be calculated. In order to compare the values thus obtained to those predicted by the Langmuir model, we rearrange expression (12.19):

$$\frac{c_{\text{inh}}}{\theta} = \frac{1}{b_{\text{L}}} + c_{\text{inh}} \tag{12.24}$$

Figure 12.32 shows a straight line, confirming that for this system the adsorption of the inhibitor follows the Langmuir isotherm.

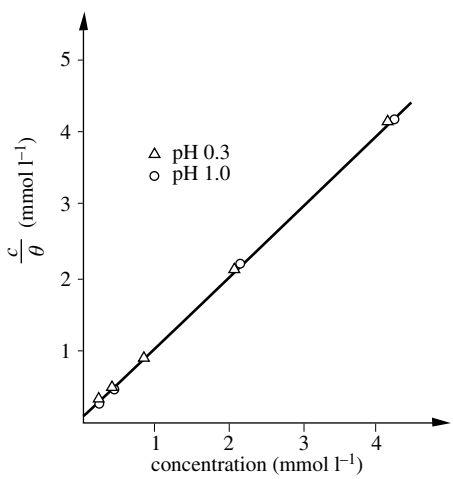


Figure 12.32 Comparison of the results given in Figure 12.31 with equation (12.24) [21].

© 2007, first edition, EPFL Press

Inhibition by formation of precipitation layers

Inhibition in acidic environments is not always due to the partial blocking of the surface by an adsorbed inhibitor. Sometimes, the inhibitor forms weakly soluble complexes with the metal ions produced by dissolution. The complexes precipitate at the surface forming a porous layer. In certain cases, the unsaturated carbon-carbon bonds of the complexes undergo polymerization reactions that render the layers more compact. The dissolution of the metal is thus slowed.

Figure 12.33 illustrates the effect of the formation of a precipitation layer on the dissolution rate: anodic polarization curves were measured in a 0.5 M H_2SO_4 solution on a copper rotating disk electrode with and without adding 25 mmol l^{-1} BTA inhibitor. In the absence of the BTA, a limiting current plateau is observed. Its height is limited by the transport of Cu^{2+} ions from the anode into the electrolyte and it varies linearly with the square root of rotation rate [22]. In the presence of BTA, the current in the plateau region is much lower and the plateau is less well pronounced, even though it still varies with the rotation rate. These observations indicate that a weakly soluble Cu-BTA complex precipitates at the surface forming a film that changes the mass transport conditions.

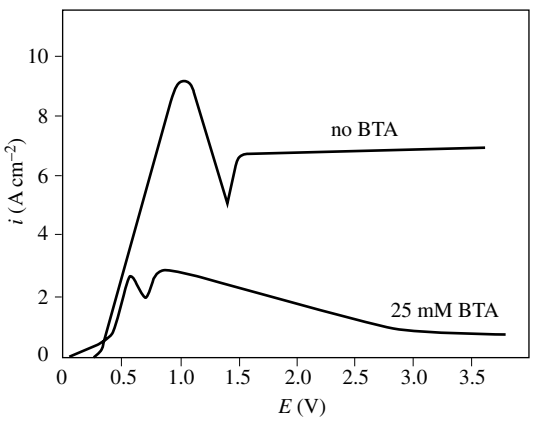


Figure 12.33 Anodic polarization curves of copper in 0.5 M H_2SO_4 in the absence and in the presence of 25 mmol l^{-1} of benzotriazole (BTA) [22].

12.3.3 Inhibitors for neutral environments

Corrosion in neutral environments is most often due to the reaction of dissolved oxygen with the metal. The risk of attack is therefore considerably reduced if one limits the access of oxygen to the metal surface (cathodic inhibition). Alternatively, one can diminish the rate of corrosion by bringing about passivation of the metal (anodic inhibition). Corrosion can also be reduced by inhibitors that adsorb onto the surface and block the reactive sites.

© 2007, first edition, EPFL Press

Inhibition by adsorption

Some inhibitors used in neutral environments protect the metal by adsorbing onto the surface as described for acid solutions. Examples include benzoates, tartrates or salycilates among other. Because neutral solutions contain a low concentration of protons the inhibitor can have an effect on the pH at the metal surface, in addition to blocking parts of the surface or changing the corrosion mechanism. Metal and oxide surfaces that are in contact with an aqueous solution are covered by adsorbed water. Any other species, inhibitors or electrolyte anions, that adsorb on the surface must replace one or several water molecules [23]. We can therefore write the adsorption reaction of an inhibitor in the following general way:

$$Inh_{aq} + x\ H_2O_{ads} \rightarrow Inh_{ads} + x\ H_2O_{aq} \tag{12.25}$$

Here *Inh* designates the inhibitor molecule, and the subscripts *aq* and *ads* designate species in the aqueous phase and adsorbed at the surface, respectively. The stoichiometric coefficient x depends on the size of the inhibitor molecule and the structure of the metal surface.

If the electrolyte contains strongly adsorbing anions, the inhibitor competes with these for adsorption sites. The adsorption behavior of inhibitors and the inhibition efficiency therefore vary with the electrolyte composition, notably with the type of anions present and their concentration. As an illustration Figure 12.34 shows polarization curves measured on iron in different electrolyte solutions of pH 7.5 in presence or not of an experimental organic inhibitor (N-ethyl-morpholine salt of w-benzoyl alcanoic acid) [24]. Figure (a) shows that the anodic inhibiting effect is greatest in distilled water and in the non adsorbing perchlorate solution and weakest in sodium sulfate. Figure (b) shows that by increasing the sodium sulfate concentration the inhibiting effect in the active potential region diminishes until it completely disappears at a sulfate concentration of 0.12 M. This is a further indication that sulfate ions compete with the inhibitor for adsorption sites.

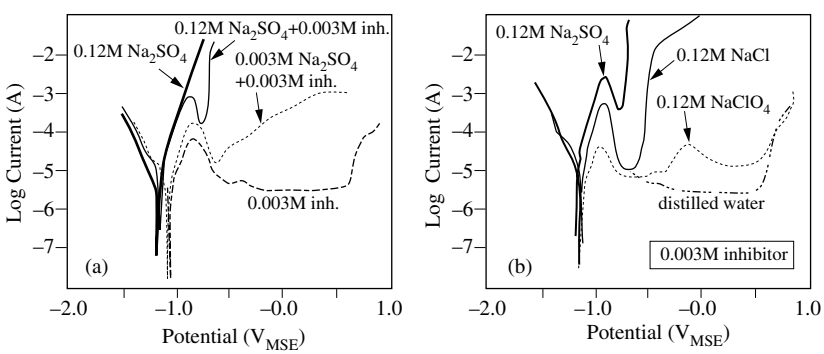


Figure 12.34 Polarization curves of iron; (a) in 0.012 M and 0.03 M sodium sulfate with and without inhibitor and 0.03 M inhibitor in water; (b) in 0.12 M sodium salt solutions of different anions in presence of 0.03 M inhibitor [24] .

© 2007, first edition, EPFL Press

A similar effect of anions on the adsorption behavior of inhibitors is also observed on oxides [24]. Figure 12.35 shows adsorption measurements on iron oxide powder performed with a inhibitor molecule similar as used for the experiments of Fig. 12.34. The behaviour in different neutral electrolyte solutions and in water is compared. It follows from these data that the inhibitor adsorbs less readily from the sodium sulfate solution than from distilled water or from perchlorate solution, the reason being that sulfate ions adsorb more strongly than perchlorate ions. The examples illustrate the importance of competitive adsorption for corrosion inhibition and they point out the complexity of the phenomena involved. While the presence of adsorbing anions affects the inhibition efficiency of adsorption type inhibitors, the presence of certain cations, for example Ca^{2+}, can have a profound effect on the formation of three-dimensional surface films and thus on the efficiency of film forming inhibitors [26].

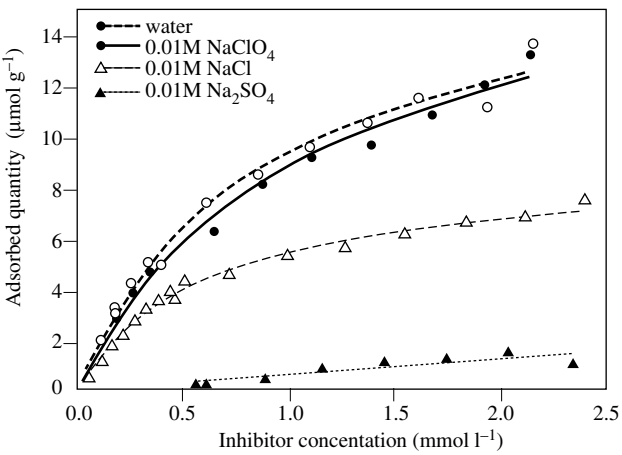


Figure 12.35 Measured adsorption isotherms of an organic inhibitor on iron oxide powder in water, 0.01 M $NaClO_4$, Na_2SO_4 and NaCl [25].

Inhibition by passivation

A typical application of this principle is the protection of steel in water-cooling circuits. Two types of inhibitors are found to be effective for this:

- oxidizing inhibitors;
- buffering agents that maintain a high pH in the vicinity of the metal surface.

Among the oxidizing inhibitors, chromates (CrO_4^{2-}) and nitrites, (NO_2^-) are most prominent. Molybdates (MoO_4^{2-}) and nitrates (NO_3^-) act in a similar way, although they are less effective oxidizing agents. Oxidizing inhibitors lead to an increase of the cathodic partial current density and thus shift the corrosion potential from the active to the passive potential region (Figure 12.36). As a consequence, the corrosion rate becomes very small.

For passivation to occur, the cathodic partial current density at the passivation potential, must exceed, in absolute value, the passivation current density. In practice, for example when protecting a cooling circuit, this condition must be fulfilled

© 2007, first edition, EPFL Press

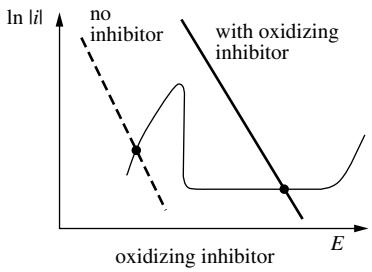


Figure 12.36 Metal passivation by an oxidizing inhibitor.

everywhere, including stagnant zones and crevices. If the metal in such locations remains in the active state, corrosion cells develop between the active and passive areas, resulting in a rapid corrosion of the active metal. Effective protection with oxidizing inhibitors requires that the inhibitor is able to reach all parts of an installation in sufficient quantity. It is therefore crucial to maintain an adequate inhibitor concentration and to control the hydrodynamic conditions.

The most powerful oxidizing inhibitors (chromates and nitrites) are toxic and their use is regulated or even prohibited by environmental legislation. Although chromates and nitrites may still be found in some closed-circuit cooling systems, they are increasingly replaced by other compounds such as molybdates or silicates.

Both the passivation current density and the passivation potential decrease with increasing pH (Chapter 6). A high pH at the metal surface therefore facilitates passivation. To maintain the pH in the neutral or alkaline range one uses ***buffering agents***. Their presence facilitates passivation and limits the corrosion rate. Buffering is particularly important in cases where corrosion cells may form because without a buffer, the pH in the vicinity of the anode decreases with time (Chapter 7). A suitably chosen buffer avoids such local acidification and thus prevents the formation of active-passive cells. To stabilize the pH in the neutral or alkaline regions, one often uses silicates, carbonates or phosphates. Small amounts of sodium hydroxide may be added to raise the pH of the solution, although NaOH itself has no buffering effect.

Inhibition by precipitation film

Many inhibitors used in neutral environments form surface layers by precipitation or polymerization reactions. The polyphosphates, $(NaPO_3)_n$ and the organophosphates belong to this category of inhibitor. Acting principally on the cathodic reaction (diffusion barrier), they lower the corrosion rate in the presence of oxygen.

Scale formation on steel in contact with natural water offers another example of inhibition by film formation due to precipitation reactions. Natural water contains many species, including the bicarbonates of magnesium and calcium. Their content in the water is expressed by the *hardness* (on the French hardness scale, a value of 1° is equivalent to 10 mg of dissolved $CaCO_3$ per liter). The hardness of natural waters is normally situated between 5° (soft water) and 25° (hard water), in the French units.

© 2007, first edition, EPFL Press

Under certain conditions, especially at elevated temperature, the dissolved bicarbonate ions are transformed into calcium carbonate according to the reaction:

$$Ca^{2+} + 2HCO_3^- \rightarrow CaCO_3 + CO_2 + H_2O \tag{12.26}$$

The carbonates of calcium and magnesium, which are only very slightly soluble, precipitate at the metal surface and form a scale. As a consequence, the oxygen accesses the metal surface less easily and the corrosion rate decreases. Scale layers also increase the ohmic resistance in the vicinity of the surface, and thus reduce the risk of galvanic corrosion (Chapter 7).

Inhibitors for cooling-water circuits

Table 12.37 provides, in qualitative fashion, the effectiveness of some commonly used inhibitors for closed cooling-water circuits [27,28]. It must be kept in mind, however, that the efficiency of an inhibitor depends on a number of factors that are not taken into account in this table: hydrodynamic conditions, chemical composition of the water, microstructure of the alloy, etc. To evaluate these effects one must turn to laboratory testing.

Table 12.37 Cooling water inhibitors for closed circuits.

	Steel	Cast iron	Zn and Zn alloys	Cu and Cu alloys	Al and Al alloys	Pb-Sn solder
chromates	+ +	+ +	+ +	+ +	+	+
nitrites	+ +	+ +	o	+ –	+	–
benzoates	+ +	–	+ –	+	+	+ +
borates	+ +	+ –	+ +	+ +	+ o	
phosphates	+ +	+ +	+	+ –	+ o	
silicates	+	+	+ +	+	+	+
tannins	+	+	+	+	+	+

+ + good protection
+ average protection
o no protection
+ o little or no protection
– increases corrosion rate
+ – positive or negative effect, depending on conditions

In practice, ***mixtures of inhibitors*** are most often used in order to combine several modes of protection. A typical example is the water-cooling circuit of an automobile. It includes a variety of metals and alloys: the motor made of aluminum alloy or cast iron, the radiator made of brass or copper and soldered joints containing of lead-tin alloys. The temperature can be below 0° C and then exceed 100° C. An adequate mix of inhibitors must be able to protect the different materials over the entire temperature range under the hydrodynamic conditions to which they are exposed during operation. In addition it should not attack the polymeric materials used in tubes and for seals.

© 2007, first edition, EPFL Press

Such a mix might contain $NaNO_3$, an oxidizing agent to protect the aluminum, borax ($Na_2B_2O_7 \bullet 5H_2O$) as a buffer, NaOH to increase the pH, MBT to protect the copper, as well as silicates and phosphates [29].

12.3.4 Inhibition of microbially influenced corrosion

Microorganisms are omnipresent in the environment: bacteria, algae, lichens, fungi etc. They play an important role for all forms of life and for the geochemical cycles of elements such as the carbon, nitrogen and sulfur [30]. They can also affect the corrosion rate of metals.

Microbially influenced corrosion

Microorganisms do not directly attack metals. Rather, their presence leads to the formation of biofilms, which modify the environmental conditions in the vicinity of the metal surface. This can cause an acceleration of the corrosion rate, hence the name ***microbially influenced corrosion*** or ***MIC*** given to this type of corrosive attack. MIC is encountered in many sectors of industrial activity, including oil drilling, water distribution and treatment, power plants, ships and civil engineering structures exposed to marine environments. In these applications microorganisms not only may accelerate corrosion, but the formation of biofilms also increases the drag at metal surfaces exposed to fluid flow (pipes, ships) and in extreme cases leads to clogging of pipes.

Biofilms

Microorganisms grow preferentially on solid surfaces where they form a biofilm. Materials of natural origin such as wood, stone and soil or of man made origin such as concrete, organic polymers and metals can serve as substrates for biofilm growth. The formation of biofilms involves several steps**.** At first, organic species such as proteins present in the environment adsorb on the surface and form a layer conducive to cell adhesion. Cells settle down on the surface and attach to the adsorbed layer. They form chemical links with the adsorbed molecules. and between themselves. The cell metabolism produces substances that together with additional adsorbing molecules lead to the formation of a polymeric material called EPS for "extra cellular polymeric substances". The cells, which become fixed in the EPS multiply and form a biofilm by incorporating additional microorganisms and solid matter from the environment. The biofilm thus grows with time and eventually can reach a substantial thickness. The biofilms are either solid or of a gel-like slimy nature. Their chemical and biological composition is complex including different organic molecules, ions, cells and solids, depending on formation conditions. To survive, microorganisms require water. Biofilms formed on solid surfaces often contain well over ninety percent of water.

Biofilms can form under aerobic or anaerobic conditions. The bacterial population differs accordingly, being constituted predominantly of aerobic or anaerobic bacteria, respectively. In certain biofilms exposed to an aerobic environment an oxygen gradient

© 2007, first edition, EPFL Press

may exist due to reaction of oxygen at the metal surface. In such a case, the outer part of the biofilm may be populated by aerobic bacteria and the inner part by anaerobic bacteria [30]. Quite generally, microorganisms change in composition and properties and adapt to changes in the environmental conditions. The detailed composition of biofilms formed on metals after prolonged exposure to biologically active media is therefore usually not well known.

Enhancement of corrosion rate by microorganisms

Microorganisms can affect the corrosion behavior of metals in several ways:

- by producing acid conditions at the surface;
- by complexing dissolved metal ions;
- by producing hydrogen sulfide;
- by forming non-homogeneous biofilms.

The metabolism of certain aerobic bacteria produces strong acids, which can accelerate corrosion. The best known example are bacteria of the *Thiobacillus* family that are able to oxidize sulfides or sulfur compounds into sulfuric acid. Certain organic acids produced by bacteria have the ability to form chelate complexes with dissolved metal ions, changing the thermodynamic conditions for corrosion (Chapter 2). In some cases the chelates may precipitate with electrolyte cations and form a film.

Most important from a corrosion view point are those microorganisms that produce sulfides by reduction of sulfate ions or other oxidized sulfur compounds. These so-called ***sulfate reducing bacteria* (*SRB*)** grow under anaerobic conditions. The sulfides formed at the metal surface can have several effects: they may interfere with the formation of a passive oxide film (Chapter 6), they may favor acid corrosion by locally creating conditions of lower pH or they may facilitate hydrogen ingress into the metal that leads to cracking (Chapter 11).

The formation of thick biofilms could in principle be beneficial if a compact film would uniformly cover a metal, preventing access of oxygen to the surface. However, biofilms usually do not form uniformly over a metal surface but usually in patches. Therefore they stimulate the formation of corrosion cells between covered and non-covered areas. The phenomenon is further enhanced if anaerobic conditions prevail at the metal surface in the areas covered by the biofilm, creating conditions for the development of anaerobic bacteria such as SRB that inhibit passivation. The following two examples illustrate the corrosive effect of microorganisms.

Corrosion of steel in the presence of sulfate reducing bacteria

Steel pipes buried in anaerobic soils such as clay sometimes exhibit a pitting type corrosion that leads to local penetration of the pipe wall. The corrosion products contain iron sulfide, which can be detected by adding a drop of hydrochloric acid: the reaction with iron sulfide releases H_2S with has the typical smell of rotten eggs. The described corrosive attack is due to the presence of SRB at the surface, which transform sulfate, a soil constituent, into sulfide. The reaction mechanism, although

© 2007, first edition, EPFL Press

not known in detail, involves the formation of corrosion cells. The population of SRB that controls the growth of the biofilm is not uniformly distributed over the surface. Sulfides favor the active state of the metal and a non-uniform distribution permits the formation of active-passive cells between sulfide rich zones and the rest of the surface. Local acidification taking place in the anodic areas further accentuates the phenomenon. The presence of SRB thus locally enhances the rate of proton reduction formation of hydrogen.

To avoid corrosion due to SRB one must create aerobic conditions. For buried pipes one generally uses a porous backfill material such as sand or gravel for this purpose. In addition, cathodic protection (see following section) can be applied. In closed systems, on the other hand, addition of biocides is an efficient method to stop bacterial growth.

Effect of aerobic bacteria on the pitting corrosion of stainless steel

Under aerobic conditions, microorganisms can induce pitting corrosion. The phenomenon is well documented for stainless steels exposed to seawater, which besides chloride ions contains a large variety of microorganisms. The presence of certain aerobic bacteria shifts the corrosion potential in the passive region to higher values and thus nearer to the critical pitting potential (Chapter 7). The immersion experiments of Figure 12.38 illustrate this behavior. In NaCl solution free of microorganisms, the corrosion potential does not vary with exposure time, but in seawater of comparable NaCl content, due to the presence of microorganisms, it increases with time. When a biocide is added to the seawater killing the microorganisms, the potential drops to the initial value [31].

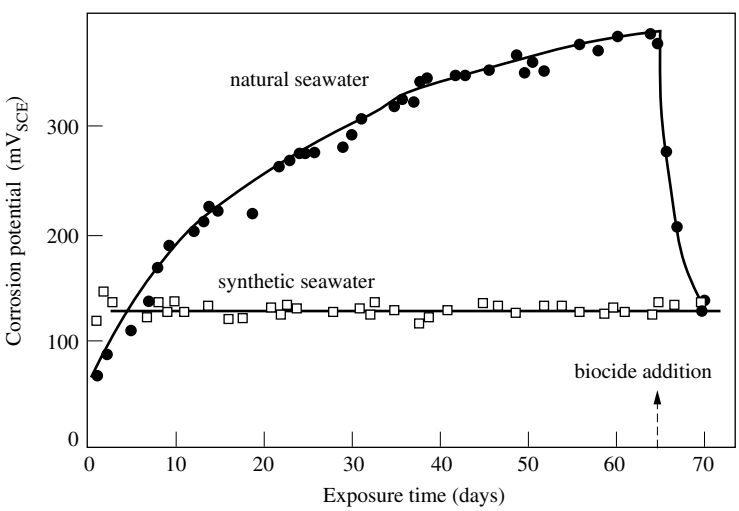


Figure 12.38 Change of corrosion potential with exposure time for 21Cr-3Mo stainless steel exposed to natural sea water containing microorganisms (•) and to artificial seawater free of microorganisms (□). After 65 days a biocide (sodium azide) was added to the sea water [31] .

© 2007, first edition, EPFL Press

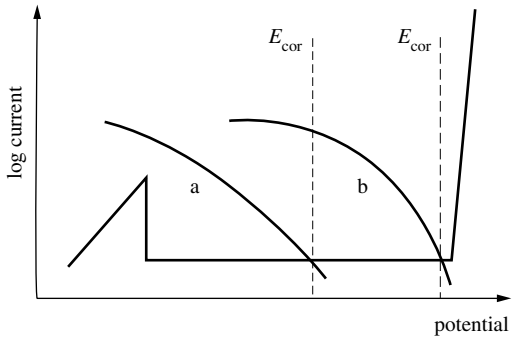


Figure 12.39 Effect of cathodic partial reaction rate on the corrosion potential of a passive metal: a higher reduction rate (curve b) leads to a higher corrosion potential.

The observed potential increase is due to an increase in the rate of the cathodic partial reaction caused by the microorganisms (Figure 12.39), in this case oxygen reduction. The mechanism by which bacteria influence the kinetics of oxygen reduction is not well known, however. It could involve a catalytic action of products of the cell metabolism, a decrease in surface pH caused by the presence of the biofilm, or the oxidation of manganese sulfide inclusions to manganese oxides acting as a catalyst [30]. Whatever the mechanism, by shifting the corrosion potential in the positive direction, the microorganisms increase the probability of pitting corrosion.

Biocides

Biocides are substances one adds to the environment in order to prevent bacterial growth. With regard to corrosion, they can be viewed as inhibitors of MIC. Certain metal ions produced by corrosion, notably copper ions, can also inhibit the growth of bacteria. However, some bacteria such as *thiobacillus* are resistant to copper ions in small concentration and for this reason, cases of MIC of copper pipes in hot water systems have been reported in the literature [32].

To control MIC in industrial installations one uses a variety of inorganic and organic biocides, a selection of which is given in Table 12.40.

Oxidizing biocides include chlorine, chlorine dioxide, bromine, ozone and hydrogen peroxide. These substances suppress the growth of bacteria, but if present in high dose they may negatively affect the corrosion behavior because their reduction at the metal surface provides an additional cathodic partial reaction. The reduction of chlorine, in addition, produces aggressive chloride ions. Bromine acts over a wider pH range than chlorine and it can be added in solid form as organic bromine compound, which facilitates handling [33]. On the other hand, bromine may pose environmental and corrosion problems. Ozone is an environment friendly biocide that can replace chlorine. It is effective already in small concentration and yields non-toxic reaction products, mostly oxygen and water.

© 2007, first edition, EPFL Press

Table 12.40 Biocides for protection against MIC (adapted from [33]).

Type	Biocide	Effective against	pH range, properties	Dose
Oxidizing	Chlorine	bacteria, algae	pH range 6.5-7.5	0.1–0.2 ppm continuous 0.5-1 pp discontinuous
	Chlorine Dioxide	bacteria, to lesser extent fungi, and algae	independent of pH;	0.1–1.0 ppm
	Bromine	bacteria, algae	broader range of pH than chlorine, less volatile	0.05–0.1 ppm
	Ozone	bacteria, biofims	no corrosive ions or other residual product formed	0.01–0.05 ppm in continuous treatment; (up to 1 ppm for biofilm removal)
	Hydrogen peroxide	bacteria	no corrosive ions or other residual products formed	50-100 ppm
Non-oxidizing	MBT*)	bacteria, algae, fungi, especially SRB	hydrolyses at pH higher than 8.0	1.5–8.0 ppm
	Isothiazolones	bacteria, algae, biofims	broad pH range desactivated by H_2S	0.9–10 ppm
	QUATS**)	Bacteria and algae	acts as surfactant mostly biodegradable	8–35 ppm
	Glutaraldehyde	bacteria, algae, fungi, biofims	broad pH range	10–70 ppm
	THPS***)	bacteria, algae and biofims	broad pH range, low environmental toxicity, dissolves iron sulfide	10–90 ppm (as active THPS)

*) Methylene-bis-thiocyanate
**) Quarternary ammonium compounds
***) Tetrakishydroxymethyl phosphonium sulfate (THPS)

Non-oxidizing biocides include a variety of organic compounds such as glutaraldehyde, quarternary ammonium salts (QUATS), isothiazolones etc. These compounds are less critical in terms of corrosion because they do not accelerate the cathodic partial reaction. Many of them are active over a wide range of pH and temperatures. THPS, a water soluble organic compound, has been found particularly effective for preventing microbial corrosion due to SRB in oilfields [33]. Biocide

© 2007, first edition, EPFL Press

formulations used in practice often contain several compounds in combination. A critical aspect of using biocides for controlling MIC is their toxicity for the environment. Before applying a biocide treatment this aspect must be thoroughly evaluated.

Other methods of protecting metals against MIC

Biocides are not the only method available to prevent and control MIC. The use of a porous backfill has already been mentioned as a simple method to prevent growth of anaerobic bacteria at the surface of buried pipes. To avoid MIC in other situations, such as cooling circuits or in marine environments, one also uses anti-fouling coatings and cathodic protection. Anti-fouling coatings are organic coatings that inhibit bacterial growth, for example by offering them poor adhesion. Cathodic protection, most often used in conjunction with coatings (see next section), is able to protect metals against general corrosion and is often effective against MIC.

Cleanliness is an important aspect of prevention of MIC in installations containing fluids [33]. Periodic cleaning can be carried out mechanically by passing sponges through a pipe or chemically, using mineral acids containing corrosion inhibitors to limit the attack of the metal surface. Quite generally, the MIC can be avoided more easily at the initial stage of bacterial growth, before thick biofilms are formed. Once formed, these may be difficult to remove and furthermore, they protect cells fixed in the EPS from contact with biocides that one adds to the fluid to inhibit MIC.

12.4 ELECTROCHEMICAL PROTECTION

12.4.1 Cathodic protection

If a sufficiently negative potential is applied to a metal its corrosion rate becomes negligibly small or zero. This forms the basis of cathodic protection, a method that is widely used to protect heavy steel structures such as drilling platforms, ships, pipelines and bridges against corrosion.

Two fundamental parameters determine the design of a cathodic protection system: the protection potential and the protection current.

Protection potential

The protection potential E_{prot}, defines the limiting value of potential at which the rate of corrosion becomes sufficiently small to be considered negligible. To protect a metal against corrosion one therefore polarizes it cathodically such that:

$$E < E_{\text{prot}} \tag{12.27}$$

© 2007, first edition, EPFL Press

The protection potential E_{prot} is defined by the following equation (Chapter 2):

$$E_{prot} = E^{\circ} + \frac{RT}{nF} \ln 10^{-6} \quad (\text{mol l}^{-1}) \tag{2.74}$$

At ambient temperature (25 °C), this expression becomes:

$$E_{prot} = E^{\circ} - \frac{0.354}{n} \tag{12.28}$$

According to this, the protection potential of steel is equal to –0.62 V, that of copper is 0.16 V and that for lead is 0.30 V.

Protection current

The protection current I_{prot} is the cathodic current one must apply to impose the protection potential. The value of the protection current is proportional to the surface area to be protected A and to the protection current density i_{prot}, which depends on electrode kinetics.

$$I_{prot} = i_{prot} A \tag{12.29}$$

Figure 12.33 shows the principle of cathodic protection for a system obeying Butler-Volmer kinetics. Because the anodic partial current density is negligible at the protection potential, i_{prot} is equal to the cathodic partial current density i_c at the potential $E = E_{prot}$. The Butler-Volmer equation thus yields for the protection current density:

$$i_{prot} = i_c = -i_{cor} \exp\left(-\frac{E_{prot} - E_{cor}}{\beta_{cH}}\right) \tag{12.30}$$

If the oxygen mass transport rather than Butler-Volmer kinetics limits the rate of the cathodic reaction at E_{prot}, the protection current density corresponds to the limiting current density of oxygen $i_{l,O2}$:

$$i_{prot} = i_{l,O2} \tag{12.31}$$

In practice, cathodic protection is most often used in combination with an ***organic coating***. The cathodic protection then serves only to prevent corrosion at sites where the coating exhibits a defect (pores, scratches, etc.). In this way the surface area of metal exposed to the environment is much smaller and the current requirements are drastically reduced compared to a non-coated metal surface. Table 12.41 indicates typical values of the cathodic current needed to protect a steel surface of one square meter, with and without an organic coating.

© 2007, first edition, EPFL Press

Table 12.41 Current required for the cathodic protection of a geometric surface of 1 m^2.

Protected surface	Environment	I_{prot} (mA)
Bare steel	soil	10 – 50
	seawater	20 – 150
Steel with a polyethylene coating of 2 mm thickness	soil	5×10^{-4}

Technological aspects of cathodic protection

Two methods of cathodic protection are commonly used in practice (Figure 12.42):

- by sacrificial anode;
- by impressed current.

A ***sacrificial anode*** cathodically protects a metal structure by forming an electrochemical cell (corrosion cell) in which the structure to be protected is the cathode. The sacrificial anode must therefore have a lower corrosion potential than the metal to be protected. To protect steel structures, magnesium, zinc, aluminum and their alloys are commonly used anode materials that fulfill this requirement. Table 12.43 compares the main characteristics of the three anode materials. Magnesium is the metal of choice for buried structures because it exhibits the greatest potential difference relative to steel and it has the best volume capacity. On the other hand, in seawater, an electrolyte with good conductivity, the potential difference is less critical and therefore aluminum and zinc are often used (ships, drilling platforms, etc.). Sacrificial anodes come in different shapes: bars, blocks, rings, etc. They are either attached directly to the object to be protected, or they are buried at a certain distance and electrically connected by wire. Figure 12.44 illustrates a set-up for the cathodic protection of a heat exchanger by a magnesium anode.

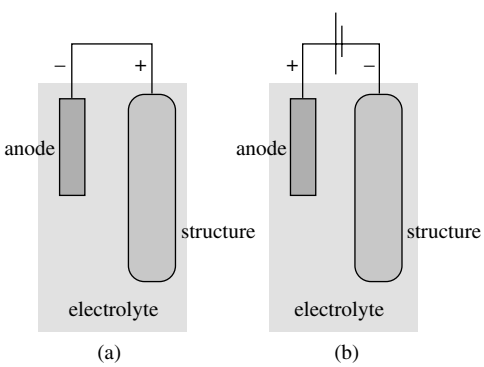


Figure 12.42 Cathodic protection: (a) by sacrificial anode; and (b) by use of an external current source.

© 2007, first edition, EPFL Press

Table 12.43 Sacrificial anodes for the protection of steel.

	Magnesium	Aluminum	Zinc
$E°$ (V)	–1.5	–1.28	–0.76
E_{cor} in soil (V)	–1.1 to –1.3	–0.6 to –0.9	–0.6 to –0.8
Potential relative to steel (V)	–0.6	–0.3	–0.2
Theoretical volume capacity (Ah/dm^3)	3840	8050	5840
Yield (%)	50	80	90
Mass needed to supply 0.1 A during 10 years (kg)	8	3.7	12

Cathodic protection by impressed current involves the use of a rectifier connected to a power line. Contrary to sacrificial anodes, which operate at a fixed potential, the use of a rectifier permits to adjust the voltage (or the current) to the particular requirements of a protection scheme. This not only allows one to optimize the electrochemical conditions for protection, but the method is also well suited to protect large surfaces. On the other hand, protection by impressed current needs more maintenance than the use of sacrificial anodes. In order to protect buried structures by impressed currents one uses consumable anodes such as scrap iron or, more often, non-consumable anodes made of iron-silicon alloy, graphite or of titanium coated with noble-metal oxides.

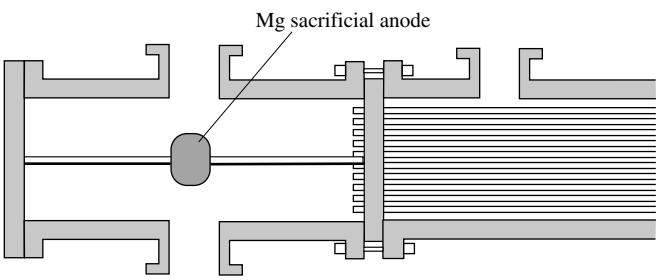


Figure 12.44 Cathodic protection of a heat exchanger by use of a magnesium sacrificial anode.

12.4.2 Electrochemical protection by passivation

A metal can be electrochemically protected from corrosion in two ways; either by applying a potential that is equal or negative to the protection potential or by imposing a potential that lies in the passive region. The second option, ***electrochemical protection by passivation,*** is based on the observation that the rate of corrosion of a metal, in the potential region of stable passivity is usually very low.

The zone of stable passivity ranges from the passivation potential E_p to the transpassivation potential (pitting potential) E_b. One can therefore reduce the corrosion of a passivating metal to a negligible level by imposing a potential E such that:

$$E_p < E < E_b \tag{12.32}$$

© 2007, first edition, EPFL Press

Depending on the value of the corrosion potential E_{cor}, this condition is achieved by applying either an anodic or a cathodic polarization to the metal.

Anodic protection

Anodic protection is applicable whenever the corrosion potential of a passivating metal is located in the active potential region: $E_{cor} < E_p$. To shift the potential of the metal into the passive region one applies an anodic polarization by means of a potentiostat (Figure 12.45). To passivate the metal, the power source must supply a current that is sufficiently high to exceed the passivation current density i_p on the entire anode area. Once the metal has been passivated, however, a much lower current density, equal to the passive current density i_{pp}, suffices to maintain the passive state. The value of i_{pp} depends on the prevailing pH and the type of metal, but is usually very small. Maintaining the anodic protection therefore requires little current, even in the absence of an organic coating. The corrosion rate of an anodically protected structure is equal to the value of i_{pp}.

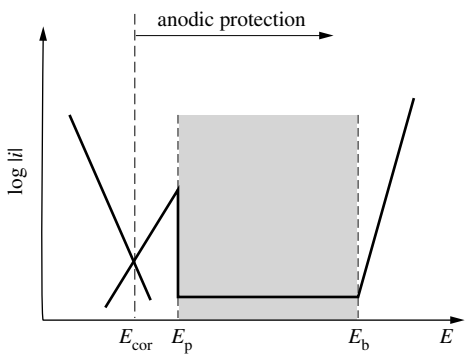


Figure 12.45 The principle of anodic protection of a passivating metal: shifting of the potential into the passive region corresponding to $E_p < E < E_b$.

Protection by cathodic polarization into the passive region

Passivating metals whose corrosion potential is located in the transpassive region (pitting corrosion, Chapter 7) can be protected by applying a cathodic polarization that shifts their potential into the region of stable passivity (Figure 12.46). The cathodic current density that is needed to maintain stable passivity by cathodic polarization depends on the kinetics of the cathodic partial reaction. In the Figure, we assume that this is the reduction of oxygen at limiting current and therefore: $i_c = i_{l,O2}$. The current density required for maintaining the passive state by cathodic polarization is normally greater than that required for anodic protection.

© 2007, first edition, EPFL Press

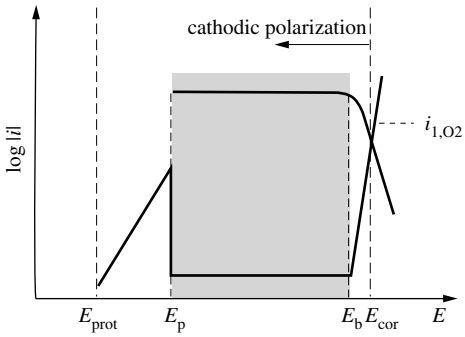


Figure 12.46 Principle of cathodic protection by passivation of a metal whose corrosion potential is located in the transpassive region: the potential shifts into the passive region corresponding to $E_p < E < E_b$. The cathodic partial current density is given by $i_c = i_{1,O2}$.

Importance of potential control

Electrochemical protection by passivation requires precise control of the potential. Indeed, unless the criterion (12.24) is satisfied over the entire surface there is a risk of accelerating locally the corrosion at the sites that are insufficiently protected. The method therefore calls for the use of a potentiostat in combination with a three-electrode set-up that includes the structure to be protected, a counter-electrode, and a reference electrode.

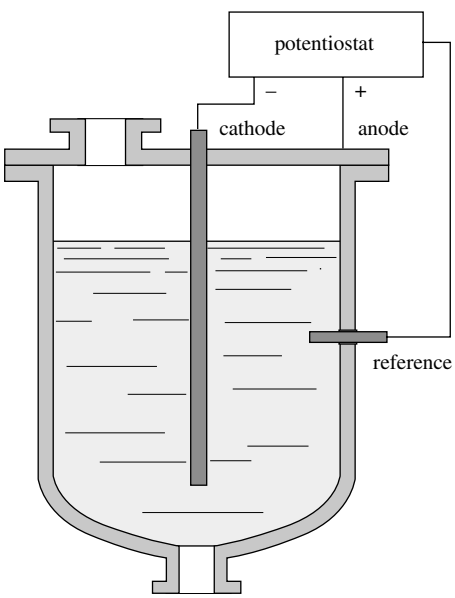


Figure 12.47 Anodic protection of a chemical reactor using a potentiostat and a reference electrode (schematic).

© 2007, first edition, EPFL Press

In practice, electrochemical protection by passivation is far less often used than cathodic protection. The method has notably been applied in the chemical industry to anodically protect chemical reactors made of steel or stainless steel, since it is well suited for non-coated surfaces. Figure 12.47 shows a schematic drawing of a three-electrode set up used for the anodic protection of the interior walls of a chemical reactor.

The geometry of the device is quite simple, with the anode and cathode positioned in concentric fashion. This facilitates potential control and avoids possible problems due to non-uniformity of the protection.

12.4.3 Electrochemical systems under ohmic control

A protection system involving a sacrificial anode is the equivalent to a galvanic corrosion cell in which, by corroding, the anode protects a metal structure that is the cathode. The driving force for current flow is the difference in corrosion potential ΔE_{cor} between the anode and the cathode (Section 7.1). The value of the current depends on the magnitude of the different resistances in the cell and is given by equation (7.9):

$$I = \frac{\Delta E_{cor}}{R_{p,I} + R_{int} + R_{ext} + R_{p,II}} \tag{7.9}$$

The resistance of the electric circuit between anode and cathode is often negligible ($R_{ext} \approx 0$). Furthermore, if the electrolyte, for example water or soil, has a low conductivity its resistance may far exceed the value of the polarization resistances $R_{p,I}$ and $R_{p,II}$. Equation (7.9) thus simplifies to (12.33) where R_Ω is the ohmic resistance of the electrolyte.

$$I = \frac{\Delta E_{cor}}{R_\Omega} \tag{12.33}$$

According to this equation, the current only depends on the difference in corrosion potential between the anode and the structure to be protected (Table 12.36) and on the ohmic resistance between the two electrodes, which is as a function of the conductivity of the electrolyte and of the geometry of the electrode arrangement.

For a protection system using an impressed current supplied by a power source, equation (12.33) is replaced by (12.34), valid when all resistances except R_Ω are negligible:

$$I = \frac{U - \Delta E_{cor}}{R_\Omega} \tag{12.34}$$

In this equation U is the output voltage of the power supply. The fact that its value can be adjusted according to needs gives more flexibility for controlling the current than the use of a sacrificial anode, whose corrosion potential in a given environment can not be varied.

© 2007, first edition, EPFL Press

Electrochemical systems in which the current flow is governed by Ohms law in the electrolyte rather than by electrode polarization are said to be under ***ohmic control***. If the electrolyte conductivity is the same everywhere the local current density on the electrodes depends only on the geometry of the set-up. For example, if two plane parallel electrodes are spaced from a central electrode at distance y_1 and y_2 the respective current densities will be inversely proportional to these distances.

Estimation of ohmic resistance in electrochemical cells

For simple cell geometries, the ohmic resistance R_Ω appearing in equations (7.9) and (12.34) can be calculated analytically. For most geometries, however, one must use numerical methods. Sometimes approximate calculations, taking into account only a part of the system, give results that are good enough for practical purposes.

To calculate the ohmic resistance for an arbitrary geometry, one must first know the potential distribution between the two electrodes. For an electrolyte of uniform composition, the potential distribution obeys the ***Laplace equation***:

$$\nabla^2 \Phi = 0 \tag{12.35}$$

The boundary conditions at the insulating walls and at the electrodes are, respectively:

$$\frac{d\Phi}{d\xi} = 0 \tag{12.36}$$

$$\frac{d\Phi}{d\xi} = -\frac{i}{\kappa} \tag{12.37}$$

Here ξ represents the direction normal to the wall and κ is the conductivity of the electrolyte.

In the following, we assume that the polarization resistance at the electrodes is negligible. The potential difference at the electrode-electrolyte interface is therefore constant, independent of current density: $\Delta\Phi = \Phi_m - \Phi_s$ = constant. To illustrate the application of Laplace's equation for these conditions, we shall calculate the ohmic resistance of cells of very simple geometry, amenable to an analytical solution.

Ohmic resistance between two plane parallel plate electrodes

In a cell with two parallel electrodes, represented in Figure 12.48, the potential gradient is one-dimensional and the Laplace equation is written as:

$$\frac{d^2\Phi}{dy^2} = 0 \tag{12.38}$$

© 2007, first edition, EPFL Press

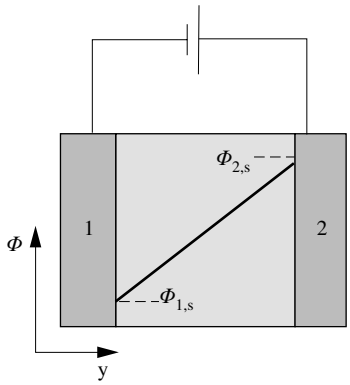


Figure 12.48 Variation of the potential in the electrolyte as a function of the distance between the two electrodes for a one-dimensional electrochemical cell.

Integration with the condition (12.37) gives :

$$\Delta\Phi_\Omega = \Phi_{2,s} - \Phi_{1,s} = -\frac{i\,L}{\kappa} \tag{12.39}$$

Here L represents the distance between the electrodes. Dividing by the current $I = i\,A$ yields the ohmic resistance :

$$R_\Omega = \left|\frac{\Delta\Phi_\Omega}{I}\right| = \frac{L}{A\,\kappa} \tag{12.40}$$

The ohmic resistance is simply proportional to the distance between the electrodes.

Ohmic resistance between concentric cylinder electrodes

The cell represented by Figure 12.49 has two concentric electrodes of radius r_1 and r_2, much smaller than their height h. The Laplace equation, expressed in cylindrical coordinates, takes on the following form:

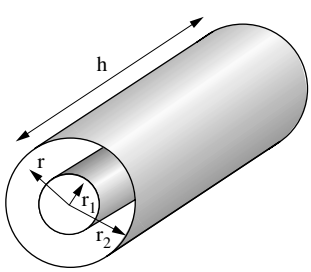


Figure 12.49 Electrochemical cell consisting of two concentric cylindrical electrodes.

© 2007, first edition, EPFL Press

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{r\,\mathrm{d}\Phi}{\mathrm{d}r}\right)=0 \tag{12.41}$$

A first integration gives: $r(\mathrm{d}\Phi/dr) = C_1$. The constant C_1 is obtained by Ohm's law:

$$i=-\kappa\frac{\mathrm{d}\Phi}{\mathrm{d}r} \tag{12.42}$$

The surface area of the cylinder is equal to $A = 2\pi\, r\, h$. With $i = I/A$ we thus obtain:

$$\frac{\mathrm{d}\Phi}{\mathrm{d}r}=-\frac{I}{2\pi\, r h\,\kappa} \tag{12.43}$$

The integration of this equation between the two potentials $\Phi_{s,1}$ and $\Phi_{s,2}$ gives:

$$\Delta\Phi_{\Omega}=\Phi_{s,2}-\Phi_{s,1}=-\frac{I}{2\pi\, h\,\kappa}\ln\frac{r_2}{r_1} \tag{12.44}$$

Dividing by I yields for the ohmic resistance:

$$R_{\Omega}=\frac{1}{2\pi\, h\,\kappa}\ln\frac{r_2}{r_1} \tag{12.45}$$

Equation (12.45) indicates that the ohmic resistance between concentric cylinders depends on the ratio of their radii, but not on their absolute values. The protection system shown in Figure 12.47 corresponds to this type of geometry with the exception of the potential distribution at the bottom, which differs.

Ohmic resistance between concentric spherical electrodes

For two concentric electrodes of spherical geometry with radii r_1 and r_2 (Figure 12.50), the Laplace equation is written as:

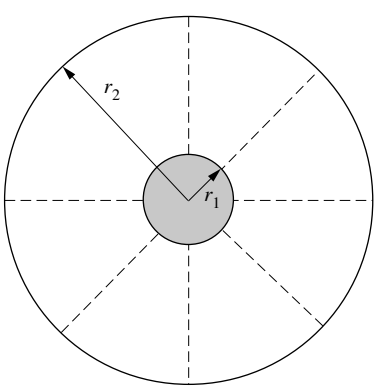


Figure 12.50 Electrochemical cell consisting of two concentric spherical electrodes.

© 2007, first edition, EPFL Press

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}\Phi}{\mathrm{d}r}\right)=0 \tag{12.46}$$

Integration yields:

$$\Delta\Phi_\Omega=\Phi_{2,s}-\Phi_{1,s}=-\frac{I}{2\pi\kappa}\left(\frac{1}{r_1}-\frac{1}{r_2}\right) \tag{12.47}$$

If the radius of the exterior sphere is much larger than that of the interior sphere ($r_2 >> r_1$) the second term can be neglected and we get:

$$\Delta\Phi_\Omega=-\frac{I}{2\pi\kappa}\frac{1}{r_1} \tag{12.48}$$

In this case, the ohmic resistance is equal to:

$$R_\Omega=\frac{1}{2\pi\kappa r_1} \tag{12.49}$$

It follows that if one of the concentric spheres is much smaller than the other the entire ohmic resistance is concentrated in the vicinity the smaller sphere. In other words, the exact size or even shape of the outer sphere plays no role. Equation (12.49) therefore is valid even if the exterior electrode is not shaped as a sphere, provided its surface is sufficiently large and located far enough from the small spherical electrode.

The described property of spherical electrodes permits to use equation (12.41) for the design of cathodic protection systems for structures buried in soil. In this case, the anode is usually much smaller than the structure it protects and often has an approximately spherical shape. Instead of having to calculate the ohmic resistance of the entire system, a more complicated task, it may be enough to estimate the ohmic resistance in the vicinity of the anode by using equation (12.49). A number of expressions for the ohmic resistance of buried anodes of other geometries are available in the literature [34].

Ohmic resistance to a disk electrode

Anodes that are imbedded in a wall often have the shape of a disk. The same geometry is found with the *rotating disk electrode*, commonly used in the laboratory for the control of mass transport conditions (Chapter 4). The ohmic resistance between a flat disk of radius r_1 embedded in an insulator, and a large counter electrode located far away, is equal to [35]:

$$R_\Omega=\frac{1}{4\kappa r_1} \tag{12.50}$$

In a similar way as for the concentric sphere, the counter electrode geometry does not enter into the final result, the disk radius being the only geometrical parameter that

© 2007, first edition, EPFL Press

determines the ohmic resistance. Equation (12.50) is often used in the laboratory to estimate the ohmic potential drop between a between a rotating disk electrode and a Luggin capillary placed far away. The disk geometry can also be used to estimate the ohmic resistance encountered in the cathodic protection of a (circular) defect in a coating where bare metal is exposed to the electrolyte.

12.4.4 Potential and current distributions on cathodically protected structures

During cathodic protection the potential and the current density on the protected structure is usually not uniform, but varies as a function of the shape of the cathode and the effective distance from the nearest anode. Figure 12.51 shows schematically the variation of the potential as a function of the distance between anodes placed in a row on a flat surface such as for example a ship hull.

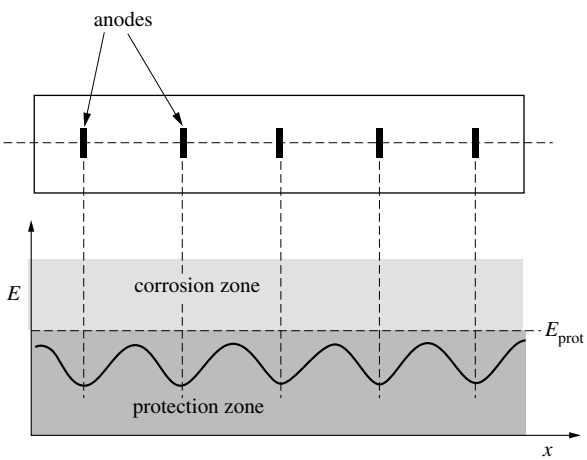


Figure 12.51 Variation of the potential as a function of the anode position during cathodic protection of a wall (schematic).

To provide cathodic protection, the entire surface must remain within a specified potential range of protection. Furthermore, the current density must exceed the protection current density at each point, but it should not become too high in areas close to an anode because this could lead to excessive hydrogen formation, which could damage the coating or create a risk of hydrogen embrittlement (Chapter 11). For the design of cathodic protection systems it is therefore important to consider the various factors that influence the current and potential distributions at an electrode.

Factors affecting the potential and current distributions on electrodes

Quite generally, the potential and the current distributions in electrochemical systems depend on a the following parameters:

© 2007, first edition, EPFL Press

- cell geometry;
- polarization resistance;
- surface films (oxides, organic coatings, etc.);
- conductivity of the electrolyte;
- mass transport conditions;
- electrode resistance.

A complete calculation of the potential and current distributions taking into account all of these parameters is seldom feasible even by numerical methods. In a first step, one therefore has to determine which of the listed parameters are most important in a given situation. This then permits to simplify the problem by neglecting the effect of the less important parameters. For example, if the polarization resistance of electrodes is small compared to the ohmic potential drops in the electrolyte, the current and potential distributions depend only on the cell geometry (*primary current distribution*). The potential distribution can then be calculated by using Laplace's equation with constant electrode potentials as a boundary condition. As an illustration, Figure 12.52 shows calculated equipotential lines and current lines at a disk electrode for primary current distribution conditions. The data show an increasingly steeper potential gradient in the vicinity of the electrode. This is consistent with the fact that the ohmic resistance between the disk and a counter electrode far away depends only on the disk radius (equation 12.50). Furthermore the data indicate that the current density on the disk electrode is non-uniform: it is much higher at the edge than in the middle.

In the above calculations the effect of electrode polarization was not taken into account. If the resistance due to electrode polarization can not be neglected, the

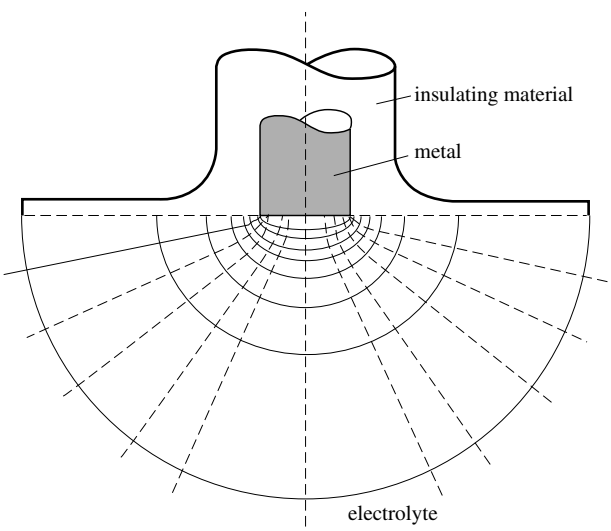


Figure 12.52 Lines of current (in dashes) and equipotential lines (solid) in the electrolyte for a rotating-disk electrode under primary current distribution conditions [35].

© 2007, first edition, EPFL Press

boundary conditions for solving the Laplace equation must be modified by including the polarization resistance or, in case of an electrode protected by a coating, the coating resistance. The described situation is that of a *secondary current distribution*. If in addition mass transport plays a critical role, the current at the electrode can no longer be computed by solving the Laplace equation only, but one must also take into account the local hydrodynamic conditions (*tertiary current distribution*). Fortunately, in cathodic protection this situation seldom arises.

Usually, the potential drop in the electrolyte is much larger than that in the electrodes and electrical leads, and the contribution of the latter can be neglected. However, there are special situations in which the electrode resistance is not negligible or may even be dominating, for example in the cathodic protection of a buried pipeline where the voltage drop in the electrolyte is relatively small because of the large volume available for current flow. The following two examples illustrate how the polarization resistance and the electrode resistance, respectively, affect the current and potential distributions on an electrode.

Current distribution in a cell of trapezoidal geometry

The calculation of the secondary current distribution in a cell of trapezoidal geometry illustrates the effect of the polarization resistance on the uniformity of the current and potential distribution at an electrode. Indeed, because the electrode potential is a unique function of the current density, $E = f(i)$, the distributions of the current and of the potential are linked.

The geometry of the model cell presented in Figure 12.53 corresponds to that of the *Hull cell*, a cell configuration widely used by electroplaters for the optimization of plating electrolytes. The Figure shows how the current density on the oblique electrode, normalized to the average current density, varies as a function of the dimensionless distance z, measured from the far end of the electrode. The results were

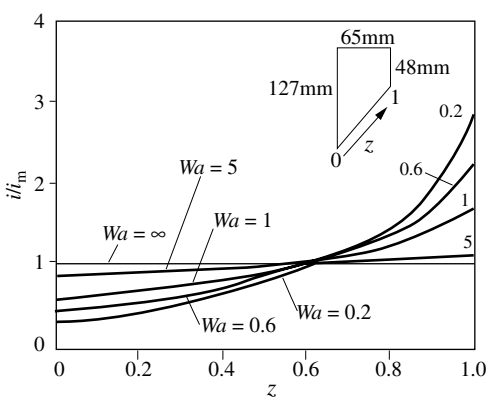


Figure 12.53 Current distribution expressed as the ratio i/i_m where i_m represents the average current density, at the oblique electrode of a trapezoidal cell (Hull cell) for different values of the Wagner number [36].

© 2007, first edition, EPFL Press

obtained by numerical solution of the Laplace equation (12.27) assuming Tafel kinetics as a boundary condition [36].

A dimensionless parameter, called the ***Wagner number***, *Wa,* characterizes the relative importance of the resistances at the interface and in the electrolyte:

$$Wa = \frac{A\left(\frac{\mathrm{d}E}{\mathrm{d}I}\right)\kappa}{L} \tag{12.51}$$

Here, A designates the geometric surface of the electrode, $\mathrm{d}E/\mathrm{d}I$ is the inverse slope of the polarization curve $I = f(E)$, κ represents the conductivity of the electrolyte, and L is a characteristic length of the system.

For an electrode reaction that obeys the Tafel equation, $\eta = -\beta_c \ln i/i_0$ (Section 4.2), the Wagner number, calculated with (12.51) for the average current density, i_{av}, is inversely proportional to the current density:

$$Wa = \frac{\beta_c \kappa}{i_{av} L} \tag{12.52}$$

On the other hand, for an electrode reaction, near its equilibrium potential and obeying linear kinetics, $i = i_0\ (1/\beta_a + 1/\beta_c)\ \eta$, the Wagner number, obtained in an analogous fashion does not depend on current density:

$$Wa = \frac{\kappa}{i_0\left(1/\beta_a + 1/\beta_c\right)L} \tag{12.53}$$

The results presented in Figure 12.53 indicate that with increasing Wagner number the current distribution becomes more uniform. Based on the definition (12.51) this can be explained in the following way. At small values of the Wagner number ($Wa \ll 1$), the current distribution is dominated by the ohmic resistance between the electrodes and it is highly non-uniform. At high Wagner number ($Wa \gg 1$) , on the other hand, the current distribution is dominated by the polarization resistance of the electrochemical reaction and as a consequence the geometrical effects are less important. With increasing Wagner number the current distribution therefore becomes increasingly uniform and for $Wa > 5$, it practically no longer varies as a function of the distance z.

The example suggests that to obtain a uniform current distribution one needs to find conditions where the resistance at the metal-electrolyte interface is much higher than that of the electrolyte in the interelectrode space. In cathodic protection this is achieved by the application of an organic coating of high resistance. In anodic protection, the passive film causes a high polarization resistance and thus favors a uniform current and potential distribution.

Potential variation on a cathodically protected pipeline

In the previous example, the ohmic resistance of the metal electrodes was negligible relative to the other resistances of the electrochemical system. When

© 2007, first edition, EPFL Press

considering cathodic protection of a buried pipeline this condition no longer applies and one must take into account the ohmic resistance of the cathode.

Figure 12.54 shows an underground pipeline, protected by two anodes, buried deeply in the ground, separated by the distance $2b$. The ohmic resistance of the electrolyte (soil) is concentrated near the anode as described in the previous section. As a consequence, the potential in the soil is uniform over the length of the pipe. Indeed, because of the very large electrolyte volume, the ionic current density far from the anode is quite small. The cathode potential, is given by the relation $E = \Phi_m - \Phi_s + constant$, where Φ_m is the potential in the metal and Φ_s that in the electrolyte just outside the electrical double layer. The value of the constant depends on the choice of the reference electrode. For the following calculation, we arbitrarily set the value of Φ_s to zero. The cathode potential the becomes:

$$E = \Phi_m + const. \tag{12.54}$$

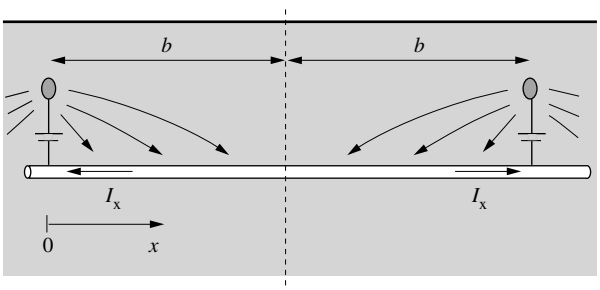


Figure 12.54 Cathodic protection of a buried pipe.

The local value of potential depends on the interfacial resistance and on that of the pipe. It can be calculated by setting up an equation for the charge balance in a small length segment dx of the pipe. Equation (12.55) states that in the segment dx the change in current flow in the pipe parallel to the surface is equal to the current flowing through the wall due to the electrochemical reaction:

$$\frac{dI_m}{dx} = -2\pi r_m i \tag{12.55}$$

In this equation I_m represents the current in the pipe, r_m is the radius of the pipe, and i_x designates the current density corresponding to the electrochemical reaction at the cylindrical surface of the pipe. The potential gradient in the pipe is obtained from Ohm's law:

$$\frac{d\Phi_m}{dx} = -\frac{I_m}{\kappa_m A_\sigma} \tag{12.56}$$

© 2007, first edition, EPFL Press

Here, κ_m designates the conductivity of the metal ($\Omega^{-1}m^{-1}$) and A_σ is the cross sectional area. The combination of the two preceding equations gives:

$$\frac{d^2\Phi_m}{dx^2} = \frac{2\pi r_m i}{\kappa_m A_\sigma} \tag{12.57}$$

For simplicity, we assume that there is a linear relation between the potential at the interface and the current density, where k_s is a constant expressed in the units $\Omega^{-1}m^{-2}$. This is a reasonable approximation for a pipe surface that is protected by an organic coating.

$$i = k_s\left(\Phi_m - \Phi_s\right) = k_s\Phi_m \tag{12.58}$$

This relation allows us to eliminate i from equation (12.57):

$$\frac{d^2\Phi_m}{dx^2} = p^2\Phi_m \tag{12.59}$$

$$p^2 = \frac{2\pi r_m k_s}{\kappa_m A_\sigma}$$

From Figure 12.54, we deduce that the potential Φ_m at any point is the sum of two terms:

$$\Phi_m = \Phi_m^{(1)} + \Phi_m^{(2)}$$

where $\Phi_m^{(1)}$ and $\Phi_m^{(2)}$ correspond to the potentials that one would measure if there were only a single anode. With the boundary conditions at $x = 0$, $\Phi_m^{(1)} = \Phi_{m,o}^{(1)}$ and at $x = \infty$, $\Phi_m^{(1)} = 0$, the solution of equation (12.59) for anode 1 takes the form:

$$\Phi_m^{(1)} = \Phi_{m,0}^{(1)} \exp(-px) \tag{12.60}$$

Similarly, for anode 2 we get:

$$\Phi_m^{(2)} = \Phi_{m,0}^{(2)} \exp(-p(2b - x)) \tag{12.61}$$

Addition of (12.60) and (12.61) gives:

$$\Phi_m = \frac{\Phi_{m(x=0)}}{\cosh(pb)} \cosh(p(x-b)) \tag{12.62}$$

Figure 12.55 shows the variation of the potentials Φ_m, Φ_1 et Φ_2 as a function of the distance from the anodes. For $x = b$, the potential is equal to:

$$\Phi_m = \frac{\Phi_{m(x=0)}}{\cosh(pb)} \tag{12.63}$$

© 2007, first edition, EPFL Press

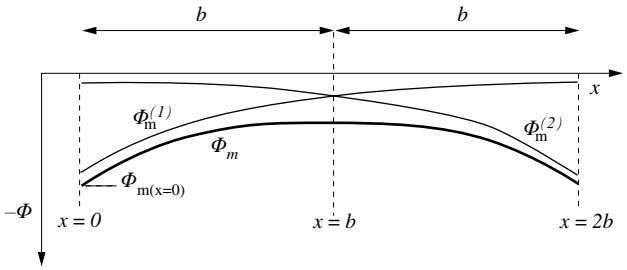


Figure 12.55 Variation of the potential $\Phi_m = \Phi_m^{(1)} + \Phi_m^{(2)}$ of the pipe as a function of the distance between the two anodes.

With equation (12.54), we can replace Φ_m in (12.63) by the potential E, measured with respect to a reference electrode:

$$E_{(x=b)} = \frac{E_{(x=0)}}{\cosh(pb)} \tag{12.64}$$

According to this, the maximum variation of the electrode potential between two anodes depends on the parameter p and hence on the ratio between the resistances in the metal and at the interface.

Spacing of anodes and current requirements

If the entire pipeline is to benefit from cathodic protection, the potential at a distance $x = b$ must be less than or equal to the protection potential: $E_{(x=b)} \leq E_{prot}$. This criterion, combined with (12.64), allows us to calculate the anode spacing for the protection of a deeply buried pipe. It will depend on the dimensionless parameter p, and on the potentials E_{prot} and $E_{(x=0)}$. The latter can be selected freely, although it should not be too negative in order to avoid significant levels of hydrogen formation.

The current required to achieve cathodic protection of the pipeline is obtained by using Ohm's law (12.65) that relates the current to the potential gradient in the metal:

$$I = 2I_{m(x=0)} = -2\kappa_m A_\sigma \left.\frac{d\Phi_m}{dx}\right|_{x=0} \tag{12.65}$$

The factor of 2 arises from the symmetry of the configuration (Figure 12.54). By differentiating (12.62) we obtain:

$$\left.\frac{d\Phi_m}{dx}\right|_{x=0} = -p\,\Phi_{m(x=0)} \tanh(pb) \tag{12.66}$$

In the absence of cathodic polarization, the potential of the pipeline corresponds to the corrosion potential. The potential of the metal at x = 0 during cathodic polarization therefore is equal to the difference

© 2007, first edition, EPFL Press

$$\Phi_{m(x=0)} = E_{(x=0)} - E_{cor} \tag{12,67}$$

With equations (12.65) and (12.66) we thus obtain:

$$I = 2\kappa_m A_\sigma \left(E_{(x=0)} - E_{cor}\right) p \tanh(pb) \tag{12.68}$$

This equation allows us to calculate the current that each anode must provide for the potential to remain equal or negative to the protection potential over the entire distance. In practice, an anode buried at a depth of several meters from the surface is able to protect a coated pipeline over a length of many kilometers [37].

Design considerations of electrochemical protection systems

Cathodic protection is widely used to protect large engineering structures in marine technology, in the petroleum industry and in civil engineering. These structures are usually of complicated shape and the placement of anodes, especially in the absence of organic coatings, requires careful consideration of the potential and current distributions. For theoretical modelling numerical methods are available such as the FEM (finite element method), BEM (boundary element method) or FDM (finite difference method). In practice, for the optimization and control of cathodic protection, experimental potential mapping is applied, using a reference electrode placed at different locations. This is particularly important in situations where interactions between different anodes and cathodes may occur. Examples are interactions in congested underground environments [38] or between the anodes used to protect the coated hull of a ship and the non-coated propeller [39].

Soils in industrial and urban areas are often densely populated with buried pipes and other metallic structures, cathodically protected or not. The electrical and electrochemical interactions between different structures can lead to the formation of corrosion cells, for example between a sacrificial anode and a copper grounding lead, and give rise to stray currents. ***Stray currents*** are *dc* currents that propagate through the ground by following a path other than the one anticipated. If a stray current enters into a buried metal structure and then leaves it at some other point, corrosion by anodic dissolution occurs there. Transport systems that use direct current (tramways, railroads) are a common source of stray currents, but cathodic protection systems can also be at their origin. To avoid corrosion, one must prevent stray currents from entering a given metal structure such as a buried pipe. This is achieved by providing an organic coating and by cathodic protection. To detect stray currents one uses potential mapping with reference electrodes. The successful application of electrochemical protection requires one to be aware of possible effects on corrosion that result from the interactions with other structures.

© 2007, first edition, EPFL Press

General Bibliography

L. L. Shreir, R. A. Jarman, and G. T. Burstein, eds., *Corrosion*, 3rd edition, vol. 2, Butterworth-Heinemann Oxford (2000).

R. Lambourne, ed., *Paint and Surface Coatings*, Ellis Horwood Ltd., Chichester (1987).

W. Baeckmann, W. Schwenk, and W. Prinz, *Handbook of Cathodic Corrosion Protection*, 3rd edition, Elsevier (1997).

References

[1] R. Baboian, ed., *NACE Corrosion Engineer's Reference Book*, 3rd Edition, NACE International, Houston Texas, 448 pp. (2002).

[2] *ASM Metals Handbook,* 9th ed., vol. 13, ASM Metals, Park Ohio, 1415 pp. (1987).

[3] G. Kreysa, M. Schütze, eds., *Corrosion Handbook*, 2nd ed., 13 vols., Wiley-VCH, Weinheim, 6500 pp. (2004-09); and 1st ed.: D. Behrens Editor, *DECHEMA Corrosion Handbook*, VCH Verlagsgesellschaft, Weinheim (1991).

[4] D. J. Renzo, *Corrosion Resistant Materials Handbook*, 4th ed., Noyes Data Corp., Park Ridge USA (1985).

[5] H. H. Uhlig, R. W. Revie, *Corrosion and Corrosion Control*, 3rd ed., John Wiley, New York (1985) p. 96.

[6] A. J. Sedriks, *Corrosion NACE* 45, 510 (1989).

[7] R. W. Schutz, D. E. Thomas, *Metals Handbook,* 9th ed., vol. 13, ASM Metals Park Ohio (1987) pp. 669–706, p.668.

[8] E. W. Skerrey, L. L. Shreir, eds., *Corrosion,* 2nd ed., vol. 2, Newnes Butterworth London (1976) pp.14–24.

[9] E. H. Hollingsworth, H. Y.Hunsicker, *Metals Handbook*, 9th ed., vol. 13, ASM Metals, Park Ohio (1987) pp. 583–609, p. 607.

[10] M. Joly, Y. Ballet, *Techniques de l'ingénieur A* (Paris) 7650 (1984) pp. 1 – 17.

[11] J. H. W. de Wit, D. H. van der Weijde, G. Ferrari, *Corrosion Mechanisms in Theory and Practice*, 2nd ed., P. Marcus, ed., Marcel Dekker Inc., New York (2002) p. 683.

[12] H. Leidheiser, Corrosion Mechanisms F. Mansfeld editor, Marcel Dekker New York 1987, p165

[13] R. T. Ruggieri, T. Beck, *Corrosion-NACE*, 39, 452 (1983).

[14] M. Favre, D. Landolt, *Corr. Sci.* 34, 1481 (1993).

[15] M. Rohwerder, M. Stratmann, P. Leblanc, G. S. Frankel, *Analytical Methods in Corrosion Science and Engineering,* P. Marcus, F. Mansfeld, eds., Taylor and Francis, Boca Raton (2005) p.605.

[16] A. Leng, H. Steckel, M. Stratmann, *Corr. Sci.* 41, 541 (1999).

[17] J. D. Crossen, J. M. Sykes, T. Zhai, G. A. D. Briggs, *Faraday Discuss.* 107, 417 (1997).

[18] G. Trabanelli, *Corrosion* 47, 410 (1991).

© 2007, first edition, EPFL Press

[19] E. McCafferty, *Corrosion Control by Coatings*, H. Leidheiser. ed., Science Press, Princeton N.J. (1979) p.279.
[20] G. Trabanelli, *Corrosion Mechanisms*, F. Mansfeld, ed. Marcel Dekker, New York (1987) p. 119.
[21] R. J. Chin, K. Nobe, *J. Electrochem. Soc.* 118, 545 (1971).
[22] R. Alkire, A. Cangellari, *J. Electrochem. Soc.* 136, 913 (1989).
[23] P. Kern, D. Landolt, *J. Electrochem. Soc.* 148, B228 (2001).
[24] P. Agarwal, D. Landolt, *Corr. Sci.* 40, 673 (1998).
[25] M. Frey, S. G. Harris, J. M. Holmes, D. A. Nation, S. Parsons, P. A. Tasker, R. E. P. Winpenny, *Chem. Eur. J.* 6, 1407 (2000).
[26] J. Telegdi, M. M. Shaglouf, A. Shaban, F. H. Karman, I. Betroti, M. Mohai, E. Kalman, *Electrochim. Acta* 46, 3791 (2001).
[27] J. Weber, *Werkstoffe Korrosion*, 26, 142 (1975).
[28] A. D. Mercer, L. L. Shreir, eds., *Corrosion* 2nd ed., vol. 2, Newnes Butterworth, London (1976) p. 18:13.
[29] C. C. Nathan, ed., *Corrosion Inhibitors*, NACE Houston, Texas (1972) p. 173.
[30] D. Thierry, W. Sand, *Corrosion Mechanisms in Theory and Practice*, 2nd ed., P. Marcus, ed., Marcel Dekker, New York (2002).
[31] V. Scotto, R. Di Cintio and G. Marcenaro, *Corr. Sci.* 25, 185 (198).
[32] W. Fischer, H. H. Paradies, D. Wagner, I. Hänssel, *Materials and Corrosion* 43, 56 (1992).
[33] H. A. Videla, "Prevention and Control of Biocorrosion," *International Biodeterioration & Biodegradation* 49(4), 259 (2002).
[34] W. Baeckmann, W. Schwenk, W. Prinz Editors, *Handbuch des kathodischen Korrosionsschutzes*, 3rd ed., Verlag Chemie, Weinheim (1989) p. 516.
[35] J. Newmann, *J. Electrochem. Soc.,* 113, 501 (1966).
[36] M. Matlosz , C. Creton, C. Clerc, D. Landolt, *J. Electrochem. Soc.* 134, 1144 (1987).
[37] D. K. Lysogorski, W. H. Hartt, *Corrosion* 60, 815 (2004).
[38] T. Huck, *Materials Performance* 44, 18 (2005).
[39] H. F. Arendt, *Materials and Corrosion*, 56, 271 (2005).

© 2007, first edition, EPFL Press

ANNEXES

A.1 CONSTANTS AND UNITS

Table A.1 Physical Constants

Name	Symbol	Value	Units
gas constant	R	8.3145	J mol^{-1} K^{-1}
Faraday's constant	F	96485	C mol^{-1}
charge of the electron	q_e	1.602×10^{-19}	C
permittivity constant	ε_0	8.95×10^{-12}	C V^{-1} m^{-1} C^2 N^{-1} m^{-2}
Plank's constant	h	6.626×10^{-34}	J s
Boltzmann's constant	k	1.38066×10^{-23}	J K^{-1}
Avogardo's number	N_L	6.022×10^{23}	atom mol^{-1}
gravitational constant	g	9.80665	m s^{-2}
mass of the electron	m_e	9.109×10^{-31}	kg
speed of light	c	2.998×10^{8}	m s^{-1}

Numerical values for 25 °C:

$$\frac{RT}{F} = 0.02568 \text{ V} \cong 0.026 \text{ V}$$

$$2.303 \frac{RT}{F} = 0.05914\, V \cong 0.059 \text{ V}$$

$$kT = 0.026 \text{ eV}$$

$$RT = 2478 \text{ J mol}^{-1}$$

© 2007, first edition, EPFL Press

Table A.2 Conversion of units.

Unit	Conversion
time duration	1 year = 3.15×10^7 s
energy	1 cal = 4.184 J
	1 J = 10^7 erg
	1 eV = 1.602×10^{-19} J atom^{-1} = 96.485 kJ mol^{-1}
	1 V A s = 1 J
force	1 N = 1 kg m = 10^5 dyn
	1 kg$_f$ = 9.80665 N
length	1 m = 10^6 μm = 10^9 nm
	1 nm = 10 Å
	1 Å = 10^{-10} m
	1 in (inch) = 2.54 cm
	1 ft (foot) = 0.3048 m
atomic mass	1 g mol^{-1} = 10^{-3} kg mol^{-1} = 1 kg kmol^{-1}
density	1 g cm^{-3} = 10^3 kg m^{-3} = 1 kg dm^{-3}
pressure	1 N m^{-2} = 1 Pa
	1 N m^{-2} = 10^6 N m^{-2} = 1MPa
	1 bar = 10^5 Pa
	1 atm = 760 torr
	1 torr = 133 Pa
	1 bar = 0.987 atm
	1 atm = 1.013 bar = 0.1013 MPa
	1 psi(lb in^{-2}) = 6.895×10^3 N m^{-2}
viscosity	1 P (poise) = 0.1 N s m^{-2}
	1 cP (centipoises) = 10^{-3} N s m^{-2}
kinematic viscosity	1 St (Stokes) = 10^4 m^2 s^{-1}

© 2007, first edition, EPFL Press

A.2 PERIODIC TABLE OF THE ELEMENTS

atomic number

element

atomic mass (g mol^{-1})

density (g cm^{-3})

1 H 1.0																	2 He 4.0
3 Li 6.9 *0.53*	4 Be 9.0 *1.85*											5 B 10.8 *2.34*	6 C 12.0 *2.26*	7 N 14.0	8 O 16.0	9 F 19.0	10 Ne 20.2
11 Na 23.0 *0.97*	12 Mg 24.3 *1.74*											13 Al 27.0 *2.70*	14 Si 28.1 *2.33*	15 P 31.0 *1.82*	16 S 32.1 *2.07*	17 Cl 35.5	18 Ar 39.9
19 K 39.1 *0.86*	20 Ca 40.1 *1.55*	21 Sc 45.0 *3.0*	22 Ti 47.9 *4.51*	23 V 50.9 *6.1*	24 Cr 52.0 *7.19*	25 Mn 54.9 *7.43*	26 Fe 55.8 *7.86*	27 Co 58.9 *8.9*	28 Ni 58.7 *8.9*	29 Cu 63.5 *8.96*	30 Zn 65.4 *7.14*	31 Ga 69.7 *5.91*	32 Ge 72.6 *5.32*	33 As 74.9 *5.72*	34 Se 79.0 *4.79*	35 Br 79.9	36 Kr 83.8
37 Rb 85.5 *1.53*	38 Sr 87.6 *2.6*	39 Y 88.9 *4.47*	40 Zr 91.2 *6.49*	41 Nb 92.9 *8.4*	42 Mo 95.9 *10.2*	43 Tc 98 *11.5*	44 Ru 101.0 *12.2*	45 Rh 102.9 *12.4*	46 Pd 106.4 *12.0*	47 Ag 107.8 *10.5*	48 Cd 112.4 *8.65*	49 In 114.8 *7.31*	50 Sn 118.7 *7.30*	51 Sb 121.8 *6.62*	52 Te 127.6 *6.24*	53 I 126.9	54 Xe 131.3
55 Cs 132.9 *1.90*	56 Ba 137.3 *3.5*	57* La 138.9 *6.17*	72 Hf 178.5 *13.1*	73 Ta 180.9 *16.6*	74 W 183.9 *19.3*	75 Re 186.2 *21.0*	76 Os 190.2 *22.6*	77 Ir 192.2 *22.5*	78 Pt 195 *21.4*	79 Au 197.0 *19.3*	80 Hg 200.6 *13.6*	81 Ti 204.4 *11.85*	82 Pb 207.2 *11.4*	83 Bi 209.0 *9.8*	84 Po 209 *9.2*	85 At 210	86 Rn 222
87 Fr 223	88 Ra 226 *5.0*	89** Ac 227															

*														
	58 Ce 140 *6.67*	59 Pr 141 *6.77*	60 Nd 144 *7.0*	61 Pm 145	62 Sm 150 *7.54*	63 Eu 151	64 Gd 157 *7.89*	65 Tb 159 *8.27*	66 Dy 162.5 *8.54*	67 Ho 165 *8.8*	68 Er 167 *9.05*	69 Tm 169 *9.33*	70 Yb 173	71 Lu 175 *9.84*

**						
	90 Th 232 *11.7*	91 Pa 231 *15.4*	92 U 238 *19.1*	93 Np 237	94 Pu 244	95 – 107

© 2007, first edition, EPFL Press

A.3 PROBLEMS

Chapter 1

1.1

Write out the stoichiometric equations for the following reactions:

- the corrosion of iron in de-aerated sulfuric acid solution;
- the corrosion of copper in aerated sulfuric acid solution;
- the corrosion of copper in aerated hydrochloric acid solution;
- the corrosion of zinc in hydrochloric acid solution.

1.2

The rate of corrosion of a common steel in de-aerated acidic solution is 30 μm year^{-1}. Calculate the anodic partial current density in μA cm^{-2}.

1.3

A 1 g aluminum sample dissolves in an acid. Calculate the volume of hydrogen produced by the reaction (25 °C, 1 bar).

1.4

The rate of hydrogen production during the corrosion of Ni in acidic conditions is measured to be 0.50 ml h^{-1} (1 bar, 25 °C). The sample has a surface of 20 cm^2. Calculate the corrosion rate in mm year^{-1}.

1.5

A Fe-13 Cr alloy (by weight %) dissolves under acidic conditions to form Fe^{2+} and Cr^{2+} ions. The anodic partial current density is 2 mA cm^{-2}. Calculate the weight loss in mg dm^{-2} day^{-1}.

1.6

Water circulates in closed circuit inside a steel heating system. The internal steel surface exposed to water is 10 m^2. To fill the system, 300 l of cold water saturated with oxygen (8 mg O_2 dm^{-3}) is required. Assuming that all the oxygen reacts with the steel, calculate the thickness of the wall lost to corrosion (in μm). Assume that the iron oxidizes to Fe_3O_4.

Chapter 2

2.1

Use the data given in Figure 2.2 to estimate the partial pressure of oxygen in equilibrium with Cu_2O at 1200 °C.

© 2007, first edition, EPFL Press

2.2

At high temperatures, carbon dioxide can react with iron as follows:

$$Fe + CO_2 = FeO + CO$$

Calculate the equilibrium partial pressure of CO_2 at 1100 °C. The total pressure $P = P_{CO} + P_{CO2}$ is 1 bar.

2.3

Calculate the equilibrium vapor pressures of water and hydrogen corresponding to the following reaction:

$$Co + H_2O = CoO + H_2$$

The temperature is 950 °C, and the total pressure $P = P_{H2} + P_{H2O}$ is equal to 1 bar.

2.4

Calculate the activity coefficient of the Fe^{2+} ion at 25 °C:
a) in a 0.01 M $FeCl_2$ solution;
b) in a solution of 0.01 M $FeCl_2$ + 0.05 M HCl.

2.5

Calculate the reversible potential of the hydrogen electrode in a solution of pH 8. The partial pressure of hydrogen is 0.5 bar and the temperature is 60 °C.

2.6

Will copper undergo corrosion in a de-aerated solution at pH 0.5 containing 10^{-6} mol l^{-1} $CuSO_4$? We assume that: $T = 25$ °C, $P_{H2} = 1$ bar.

2.7

In aqueous solution, ferric ions are in equilibrium with hydrolyzed $FeOH^{2+}$ and $Fe(OH)_2^+$ ions:

$$Fe^{3+} + H_2O = FeOH^{2+} + H^+$$
$$FeOH^{2+} + H_2O = Fe(OH)_2^+ + H^+$$

Calculate the concentrations of the Fe^{3+}, $FeOH^{2+}$, and $Fe(OH)_2^+$ ions in an electrolyte of pH 4. The total concentration of trivalent iron is equal to 0.1 mol l^{-1}, the temperature is 25 °C. The table below gives the free energies of formation of the different ionic species:

© 2007, first edition, EPFL Press

Species	Fe^{3+} (aq)	$FeOH^{2+}$ (aq)	$Fe(OH)_2^+$ (aq)
$\Delta G°$ (kJ mol^{-1})	–4.6	–229.4	–438.1

Remark: To solve this problem, activities are to be replaced by concentrations.

2.8

Calculate the complexation constant for silver in a cyanide-containing environment (25 °C) corresponding to the following equilibrium:

$$Ag^+ + 2\ CN^- = Ag(CN)_2^-$$

2.9

The equilibrium potential of the cell:

$$Pt \mid H_{2(1\ bar)} \mid HCl_{(0.12\ M)} \mid AgCl \mid Ag$$

is equal to 0.342 V (25 °C). The hydrogen partial pressure is equal to 1 bar and the concentration of hydrochloric acid is 0.12 mol l^{-1}. Calculate the activity coefficient of the hydrochloric acid.

2.10

Find the reversible potential of the following cell:

$$Pt' \mid O_{2(0.002\ bar)} \mid KOH_{(0.01\ M)} \mid O_{2(0.2\ bar)} \mid Pt''$$

The temperature is 80 °C. Identify the positive electrode and the anode.

2.11

Calculate the protection potential of nickel with respect to the saturated calomel electrode at a temperature of 25 °C.

2.12

Calculate the maximum dissolution rate of copper (in mol m^{-2} s^{-1})in 1 M sulfuric acid. The potential of the metal at 25 °C is 0.160 V. The mass transport coefficient is $k_L = 10^{-5}$ m s^{-1}. Does the dissolution rate depend on the acid concentration?

2.13

Calculate the equilibrium potential in an environment of pH 8 of the cell:

$$Pt \mid H_{2(1\ bar)} \mid Ni^{2+}, OH^-, H_2O \mid Ni(OH)_{2(s)} \mid Ni$$

The solubility constant of $Ni(OH)_2$ is equal to $K = 1.6 \times 10^{-16}$ mol^2 l^{-2}.

© 2007, first edition, EPFL Press

2.14

The standard potentials of the following electrode reactions are known:

$$Cu(NH_3)_2^+ + e = Cu + 2\ NH_3 \qquad E° = -0.10\ V$$
$$Cu(NH_3)_4^{2+} + e = Cu(NH_3)_2^+ + 2\ NH_3 \qquad E° = 0.10\ V$$

a) Calculate the standard potential of the reaction

$$Cu(NH_3)_4^{2+} + 2\ e = Cu + 4\ NH_3$$

b) Indicate the predominant complex species formed during corrosion of copper in ammonia solution and calculate the protection potential of copper in 0.1 M NH_3 at a temperature of 25 °C.
c) Calculate the complexion constant of Cu^+ and of Cu^{2+} in ammonia solution corresponding to the reactions:

$$Cu^+ + 2\ NH_3 = Cu(NH_3)_2^+$$
$$Cu^{2+} + 4\ NH_3 = Cu(NH_3)_4{}^{2+}$$

2.15

The standard potentials of the following electrode reactions are known:

$$Cr^{2+} + 2e = Cr \qquad E° = -0.90\ V$$
$$Cr^{3+} + 3e = Cr \qquad E° = -0.74\ V$$

a) Will the anodic dissolution of chromium produce mostly Cr^{2+} ions or Cr^{3+} ions?
b) Show that the mole fraction of Cr^{3+} ions in equilibrium with metallic chromium varies with the total chromium ion concentration according to the expression :

$$X_{Cr3+} = 7.32 \times 10^{-9}\ (1 - X_{Cr3+})^{3/2} c^{1/2}$$

where X_{Cr3+} is the mole fraction of the Cr^{3+} ions defined as $X_{Cr3+} = c_{Cr3+}/c$, where the total chromium ion concentration is given by $c = c_{Cr2+} + c_{Cr3+}$.
c) Calculate the concentrations Cr^{3+} and Cr^{2+} ions in equilibrium with metallic chromium for a total chromium ion concentration of $c = 0.01$ mol l^{-1}.

Chapter 3

3.1

At 25°C the equilibrium vapor pressure of water is equal to 0.023 bar. Calculate the vapor pressure of a water droplet of 1 μm diameter for these conditions.

3.2

What pressure needs to be applied in order to force water through a frit with pore diameter of 3 μm?

© 2007, first edition, EPFL Press

3.3

Calculate the number of atoms per square meter of a monocrystalline surface of copper for the (111) and (100) orientations. The atomic radius of copper is 0.1278 nm.

3.4

In an high-vacuum chamber, the partial pressure of carbon monoxide is $P_{CO} = 10^{-5}$ Pa (25 °C). By assuming a sticking coefficient of 0.5, calculate the time needed to form an monolayer of adsorbed CO. The number of adsorption sites is 10^{19} m^{-2}.

3.5

Use the results given in Figure 3.12 in order to determine the heat of adsorption of CO, corresponding to a coverage of $\theta = 0.5$, on a (111) surface of a platinum monocrystal.

3.6

The composition of an oxide film formed on an iron-chromium alloy is studied by XPS. The thickness of the film is equal to 2 nm, and it is covered with a monolayer of carbon because of contamination. The atomic radius of carbon is equal to 0.09 nm, the escape depth of photo-electrons for trivalent chromium ions in both the oxide and in the contamination layer is 1.9 nm. Calculate the attenuation of the Cr^{3+} signal by the carbon film.

3.7

An iron-nickel alloy is analyzed by ISS with the use of helium ions as the primary beam. The angle θ between the incident and reflected ions is 90° and the primary energy is 1 KeV. What does the resolution of the instrument (with respect to energy) have to be if we want to separate the iron and nickel peaks?

3.8

Calculate, according to the Gouy-Chapman model, the diffuse double layer capacity (in μF cm^{-2}) in a 0.01 M NaCl solution. The difference between the applied potential and the potential of zero-charge corresponding to $\Delta\Phi_{GC}$ is equal to 0.05 V (25 °C).

3.9

The double layer capacity of a semi-conducting electrode is measured as a function of the applied potential. It is found that:

© 2007, first edition, EPFL Press

E (V)	–0.1	0	0.1	0.2	0.3	0.5	0.7
C (μF cm^{-2})	1.4	0.77	0.60	0.50	0.44	0.36	0.32

Calculate the charge carrier density. The dielectric constant is $\varepsilon = 8.0$.

Chapter 4

4.1

The exchange current density of the reaction

$$2H^+ + 2\,e = H_2$$

on a metal electrode immersed in an acidic solution of pH 3.0 is equal to $i_0 = 1 \times 10^{-4}$ A cm^{-2} (25 °C). The cathodic Tafel coefficient is equal to $\beta_c = 55$ mV. Calculate the cathodic current density for an applied potential of $E = -600$ mV, measured with respect to the saturated calomel electrode.

4.2

The exchange current density of the reaction

$$Fe^{3+} + e = Fe^{2+}$$

in an electrolyte that contains equal amounts (0.001 mol l^{-1}) of Fe^{2+} and Fe^{3+} ions is equal to $i_0 = 5 \times 10^{-3}$ A cm^{-2}. The value of the cathodic Tafel coefficient is $\beta_c = 0.05$ V. We then add 0.019 mol l^{-1} of Fe^{3+} ions to the solution. What is the new value of the exchange current density?

4.3

The cathodic polarization curve for proton reduction in a solution of pH = 1.0 is measured using a saturated calomel electrode as the reference electrode. The working electrode has a surface of 6 cm^2, and the temperature is 25 °C. We find the following results:

$E_{(SCE)}$ (mV)	I (A)
–360	–0.0003
–480	–0.0029
–600	–0.0310
–720	–0.301

Calculate the cathodic Tafel coefficient, β_c, and the exchange current density, i_0.

© 2007, first edition, EPFL Press

4.4

In order to determine the polarization resistance of a carbon steel in acidic solution, a potential is applied and the current measured. We obtain the following values:

$E - E_{corr}$ (mV)	i (mA cm^{-2})
–15	–8.5
–10	–6.0
–5	–2.6
–2	–1.0
0	0
2	1.1
5	2.8
10	5.0
15	9.0

The anodic and cathodic Tafel coefficients are equal to: β_a = 30 mV, β_c = 50 mV. Calculate (a) the polarization resistance at the corrosion potential and (b) the corrosion current density.

4.5

The cathodic polarization curve of a nickel electrode is measured in a de-aerated acid solution. The saturated calomel electrode is used as the reference. The working electrode has a surface of 2 cm^2. The following results are obtained:

$E_{(SCE)}$ (V)	–0.55	–0.64	–0.69	–0.71	–0.73	–0.77
I (mA)	0	0.794	3.05	4.90	8.10	20.0

Calculate the corrosion current density as well as the rate of corrosion (in mm per year).

4.6

Calculate the corrosion current density and the value of the corrosion potential of a steel sample in a de-aerated 4% NaCl solution at pH 1.0 containing 10^{-2} mol l^{-1} of $FeCl_2$. The following is known:

$i_{0,Fe} = 10^{-8}$ A cm^{-2}
$i_{0,H} = 10^{-7}$ A cm^{-2}
$\beta_{a,Fe}$ = 17.4 mV
$\beta_{c,H}$ = 52 mV

© 2007, first edition, EPFL Press

4.7

The potential of a steel sample is measured in a de-aerated solution of pH 4 as a function of applied current. It is found that the slope of the cathodic Tafel line corresponds to:

$$b_{c,H} = -\frac{dE}{d_{10}\log|i|} = 120 \text{ mV}$$

The corrosion potential, relative to the saturated calomel electrode, is E_{cor}(SCE) = –0.520 V. For a cathodic current density of 1.0 mA cm^{-2}, a potential of $E_{(sce)}$ = –0.900 V is measured. Calculate the corrosion current density and the exchange current density of the cathodic partial reaction.

4.8

The corrosion potential of a zinc electrode in a de-aerated solution at pH 3.0 is equal to E_{cor} = 500 mV and the corrosion current density is i_{cor} = 10^{-6} A cm^{-2}. Upon addition of hydrochloric acid, the pH decreases to 1.0 and the corrosion current density increases to i_{cor} = 10^{-5} A cm^{-2}. What is the new corrosion potential? For this calculation, we assume that the cathodic reaction is first order with respect to H^+ ions. The cathodic Tafel coefficient is equal to β_c = 0.053 V.

4.9

Calculate the limiting current for the reduction of oxygen on a rotating disk electrode in aqueous solution for a rotation rate of 1200 rpm. The diffusion coefficient for oxygen in water is equal to 2.51×10^{-5} cm^2 s^{-1}. The oxygen concentration for a solution in equilibrium with the atmosphere is equal to 8 mg l^{-1}, and the kinetic viscosity is 0.01 cm^2 s^{-1}. What is the thickness of the diffusion layer?

4.10

Aerated water flows down a tube with an inside diameter of 19 mm. The Reynolds number is 42,000. The corrosion rate is controlled by the rate of mass transport of oxygen, whose concentration in the water is equal to 7 mg l^{-1}. The kinetic viscosity of water is equal to 0.01 cm^2 s^{-1}. Calculate the corrosion current density of the metal walls of the tube.

4.11

Calculate the corrosion rate (in mm per year) of a magnesium cylinder of diameter 4 cm turning at 4000 rpm in a solution of 0.01 M HCl with 0.1 M NaCl. The corrosion rate is limited by the diffusion of protons towards the cylinder surface. The kinetic viscosity of the solution is $\nu = 10^{-6}$ m^{-2} s^{-1}.

© 2007, first edition, EPFL Press

4.12

The idea is put forward to store highly radioactive waste in cylindrical copper containers covered with bentonite (an aluminosilicate) to be deposited into granite stone. The bentonite creates a barrier against the incursion of ground water, and, in addition, it maintains a high pH at the surface of the barrel. The thickness of the bentonite layer is 2 m. The outside of the layer is in contact with salt water that contains 5.4 ppm O_2. The diffusion coefficient of oxygen in bentonite is equal to $D = 8 \times 10^{-11}$ m^2 s^{-1}. Calculate the thickness of the wall corroded after 100,000 years, the planned storage time. For this calculation, we neglect the effect of the cylindrical geometry of the recipient.

4.13

Oxygen is reduced on a platinum rotating disk electrode at limiting current:

$$O_2 + 2\,H_2O + 4\,e \rightarrow 4\,OH^-$$

The electrolyte is a NaCl solution at pH 7.0. The oxygen concentration is 8.0 mg l^{-1}, and the rotation rate is 100 rad s^{-1}. Estimate the pH at the surface of the cathode. Does its value depend on the rotation rate?

4.14

The saturation concentration of $FeCl_2$ in water is equal to 4.25 mol l^{-1}. Calculate the anodic limiting current density for iron dissolution from a rotating disk electrode in a binary electrolyte of 2.0 mol l^{-1} $FeCl_2$ at a rotation rate of 200 rpm. The temperature is 25 °C.

4.15

Using the values found in Table 4.19, estimate the conductivity of drinking water (25 °C) that contains 1 mg l^{-1} of NaCl and 15 mg l^{-1} of $Ca(HCO_3)_2$.

4.16

A spherical ion moves in a unidirectional electric field $d\Phi/dy$. The electric force F_e acting on the ion corresponds to:

$$F_e = z_i\, q\, \frac{d\Phi}{dy}$$

with $q = 1.60 \times 10^{-16}$ C. The frictional force that slows down the ion velocity is given by Stokes law:

$$F_r = 6\,\pi\, r_i\, \eta\, \upsilon_i$$

© 2007, first edition, EPFL Press

In this equation, υ_i is the displacement velocity of the ion, r_i is its radius and η is the viscosity of the solvent (for water at 25 °C, $\eta = 10^{-3}$ kg ms^{-1}). The ion reaches a steady-state velocity when $F_e = F_r$. By using the data of Table 4.19, calculate the effective ionic radius of hydrated Fe^{2+} and Fe^{3+} ions. Why is the effective ionic radius of the Fe^{3+} ion the higher of the two?

Chapter 5

5.1

A solution in contact with a platinum electrode contains 10^{-2} mol l^{-1} of Fe^{2+} ions and 10^{-2} mol l^{-1} of Fe^{3+} ions. A potential of 0.595 V is applied to the electrode. What is the value of the measured cathodic current? Is the reaction rate under these conditions controlled by charge transfer or by mass transport? We know the following: $i_0 = 5 \times 10^{-3}$ A cm^{-2}, $\beta_c = 0.05$ V, $\delta = 10^{-3}$ cm, $D_{Fe3+} = 0.61 \times 10^{-9}$ m^{2}s^{-1}, $T = 25$ °C.

5.2

For the dissolution of iron according to the Bockris mechanism, find the values of the following quantities:

$$\beta_a \ , \quad \left(\frac{d \ln i}{d \ln c_{OH^-}}\right)_{c_{Fe2+}} , \quad \left(\frac{d \ln i_0}{d \ln c_{Fe2+}}\right)_{c_{OH-}} , \quad \left(\frac{dE}{d \ln c_{OH-}}\right)_{i, c_{Fe2+}}$$

5.3

Protons are reduced according to overall stoichiometry:

$$2H^+ + 2\,e \rightarrow H_2$$

The reaction kinetics on a given metal has been found to follow the Volmer-Tafel mechanism with the Tafel reaction being rate-limiting. Show that if the surface coverage of adsorbed hydrogen remains low, the Tafel slope is equal to $b_c = 30$ mV (25 °C, base–10 logarithm).

5.4

The impedance of a metal electrode of 2 cm^2 is measured at the equilibrium potential for the reaction (M represents a metal):

$$M^{2+} + 2\,e \rightarrow M$$

The following values are found:

© 2007, first edition, EPFL Press

Frequency (kHz)	Z_{Re} (Ω)	Z_{Im} (Ω)
20	0.03	0.50
10	0.12	0.98
5	0.47	1.88
2.5	1.60	3.20
1.5	3.28	3.93
1.25	4.00	4.00
1	4.88	3.90
0.75	5.88	3.53
0.5	6.90	2.76
0.25	7.69	1.54
0.1	7.95	0.64
0.05	7.99	0.32
0.01	8.00	0.06

Calculate the exchange current density of the electrode reaction and the double layer capacity.

5.5

The impedance of a metal electrode with a surface area of 5 cm^2 is measured. The metal dissolves in the Tafel region at a current density of 5 mA according to the reaction:

$$M \rightarrow M^+ + e$$

The following values are found for the impedance:

Frequency (kHz)	Z_{Re} (Ω)	Z_{Im} (Ω)
20’000	2.00	0.10
10’000	2.01	0.20
5000	2.03	0.41
1000	2.60	1.80
500	3.85	2.77
200	6.41	2.65
100	7.50	1.65
50	7.87	0.180.88
10	7.99	0.020.18
1	8.00	0.02

Calculate (a) the ohmic resistance between the Luggin capillary and the working electrode, (b) the charge transfer resistance, and (c) the value of the Tafel coefficient β_a.

© 2007, first edition, EPFL Press

5.6

The following mechanism is proposed for nickel dissolution:

$$Ni + H_2O \rightarrow NiOH_{ads} + H^+ + e$$
$$NiOH_{ads} \rightarrow NiOH^+ + e$$
$$NiOH^+ + H^+ \rightarrow Ni^{2+} + H_2O$$

In this mechanism, the second step is rate limiting, and the adsorption of $NiOH_{ads}$ follows the Langmuir isotherm. Find the values for the following quantities expressed in base-10 logarithms:

$$\left(\frac{d\log i_0}{d\,pH}\right)_{c_{Ni2+}}, \quad \left(\frac{d\log i_0}{d\log c_{Ni2+}}\right)_{pH}$$

5.7

The dissolution of copper in sulfuric acid involves two consecutive steps:

$$Cu \rightarrow Cu^+ + e$$
$$Cu^+ \rightarrow Cu^{2+} + e$$

Show that the reaction obeys the expression:

$$i = 2\,i_{0,II}\,\frac{\exp((1+\alpha)f\eta) - \exp(-(1-\alpha)f\eta)}{1 + \dfrac{i_{0,II}}{i_{0,I}}\exp(f\eta)} \quad \text{with } f = F/RT$$

Calculate the anodic and cathodic Tafel coefficients for the special case where the exchange current density of the first step, $i_{0,I}$ is significantly larger than that of the second step $i_{0,II}$: $i_{0,I} >> i_{0,II}$. For this calculation, we assume that the transfer coefficients for the two steps are equal: $\alpha_I = \alpha_{II} = \alpha$. In addition, the concentration of the Cu^{2+} ions at the electrode surface remains constant.

Chapter 6

6.1

By using the values of the Gibbs free energy of formation from Tables 2.9 and 6.8, calculate the standard potential of the following half-cell reaction:

$$Fe_2O_3 + 6\,H^+ + 6\,e = 2\,Fe + 3\,H_2O$$

Compare the results obtained with the passivation potential of iron as given in Figure 6.9.

6.2

Calculate the standard potential for the reduction of hematite to magnetite:

$$3\,Fe_2O_3 + 2\,H^+ + 2\,e = 2\,Fe_3O_4 + H_2O$$

© 2007, first edition, EPFL Press

6.3

The active dissolution of a passivating alloy obeys the Volmer-Butler law. The passivation potential (25 °C) depends on the pH according to the following expression:

$$E_p = 0.15 - 0.059 \text{ pH}$$

The corrosion potential and the corrosion current density of the alloy in a solution of pH 0.2 are equal to: $E_{cor} = -50$ mV, $i_{cor} = 0.10$ mA cm^{-2}. The anodic Tafel coefficient is $\beta_a = 30$ mV. Calculate the passivation current density.

6.4

The chemical passivation of a rotating stainless steel cylinder immersed in an acidic solution is studied as Fe^{3+} ions are gradually added to the solution. The Fe^{3+} ions are reduced to Fe^{2+} at the limiting current as the metal corrodes. The radius of the cylinder is 2 cm, and the rotation rate is 1000 rpm. The diffusion coefficient of the Fe^{3+} ions is equal to 0.7×10^{-5} cm^{2}s^{-1}, the cinematic viscosity of the solution is $\nu = 0.011$ cm^{2}s^{-1}. In the absence of Fe^{3+}, the anodic polarization curve indicates a passivation current density of $i_p = 30$ mA cm^{-2}, independent of the rotation rate. Calculate the concentration of Fe^{3+} required to passivate the alloy.

6.5

The passivation current density of iron decreases with increasing pH. On the basis of the experimental data given in Figure 6.16, set up an equation that describes the variation of the passivation current density of iron as a function of pH in a phosphate-containing solution. Indicate the minimum pH that still permits the spontaneous passivation of iron in a solution that contains 0.5 ppm of dissolved oxygen (1 ppm = 1 mg l^{-1}). The rate of oxygen reduction is limited by mass transport and the thickness of the diffusion layer is equal to 2.0×10^{-2} cm.

6.6

In Figure 6.15 the passivation current density of iron in 1 M sulfuric acid is shown for different rotation rates of a rotating disk electrode .
a) Give an expression for the variation of the passivation current density as a function of the square root of the rotation rate: $i_p = f(\omega^{1/2})$.
b) Estimate the Fe^{2+} ion concentration at the surface of the electrode. The diffusion constant for Fe^{2+} ions is $D_{Fe2+} = 0.5 \times 10^{-5}$ cm^{2}s^{-1} and the cinematic viscosity of the electrolyte is $\nu = 0.01$ cm^2 s^{-1}.

6.7

For an electrolyte at pH 5.0, calculate the equilibrium potential of the following reaction:

$$2\ CrO_{3(s)} + 6\ H^+ + 6\ e = Cr_2O_{3(s)} + 3\ H_2O$$

© 2007, first edition, EPFL Press

6.8

The transpassive dissolution of a stainless steel rotating disc electrode in concentrated acid reveals an anodic limiting current that varies with the rotation rate:

ω (rad s^{-1})	600	400	200	100	50	20	10
i_l (A cm^{-2})	0.262	0.25	0.227	0.20	0.172	0.134	0.107

Calculate the dissolution current density at infinite rotation rate.

6.9

The transpassive dissolution of chromium produces different hexavalent species: chromic acid (H_2CrO_4), chromate ions (CrO_4^{2-}), hydrogen chromate ions ($HCrO_4^-$) and bichromate ions ($Cr_2O_7^{2-}$). The relative concentrations of these species are determined by the following equilibria:

$$H_2CrO_4 = HCrO_4^- + H^+ \qquad K_1 = 6.3$$
$$HCrO_4^- = CrO_4^{2-} + H^+ \qquad K_2 = 3.3\times10^{-7}$$
$$2\ HCrO_4^- = Cr_2O_7^{2-} + H_2O \qquad K_3 = 33.9$$

Calculate the concentration of each species in a solution of pH 3 for the total hexavalent chromium concentration of 0.1 mol l^{-1}. For this calculation, activities are replaced by concentrations.

Chapter 7

7.1

Certain parts of a water heater made of copper are maintained at a temperature of 95 °C, others at 65 °C. The heater holds water with 0.1 ppm of Cu^{2+} ions. Estimate the theoretical voltage between the parts at different temperature. Identify the place where corrosion could potentially take place. For this calculation, we assume that the standard potential does not vary between the two temperatures of interest.

7.2

The voltage at the leads of a battery at open circuit is equal to $U_{I=0}$ = 2 V. The potentials of the anode and the cathode obey the following expressions:

$$\zeta_a = (E - E_{cor})_{anode} = 0.18 + 0.12 \log |I| \quad \text{(V)}$$
$$\zeta_c = (E - E_{cor})_{cathode} = -0.20 - 0.09 \log |I| \quad \text{(V)}$$

The internal resistance of the battery is equal to R_{int} = 0.14 Ω. Calculate the voltage at the leads when the battery supplies a current of (a): I = 0.1 A , and of (b): I = 1.0 A.

© 2007, first edition, EPFL Press

7.3

A 25 dm^2 section of sheet metal, with 5 rivets (0.5 cm^2 each), stays immersed in water that contains dissolved oxygen. The rate of oxygen reduction is the same for the entire piece and is equal to $N_{O2} = 4.0 \times 10^{-3}$ mol cm^{-2} s^{-1}.
a) The rivets are made of steel, the sheet metal is copper. Calculate the rate of corrosion of the rivets in mm per year.
b) The sheet metal is made of steel, and the rivets are copper. Calculate the average rate of corrosion of the sheet metal in mm per year.
For these calculations, we assume that the iron oxidizes into Fe^{2+} and that the corrosion of the copper is negligible.

7.4

Two 10 cm^2 sections of sheet metal, one of zinc and the other of copper, are immersed in a solution of aerated Na_2SO_4. Oxygen is reduced at the limiting current on both metals, and the limiting current density is identical: $i_{1,O2}$ = –0.05 A cm^{-2}. The corrosion potential of zinc is equal to: $E_{cor,Zn}$ = –765 mV, and that of copper is $E_{cor,Cu}$ = +10 mV. The Tafel coefficient for Zn dissolution is equal to $\beta_{a,Zn}$ = 40 mV. The resistance of the electrolyte between the two metals is equal to R_{int} = 1 Ω.
By short-circuiting the two metals (R_{ext} = 0), a corrosion cell is formed. The corrosion of the copper stops, while that of the zinc increases. Calculate the following values for the cell:
a) the corrosion rate of the zinc (in A cm^{-2});
b) the potential of the zinc;
c) the electric current between the cathode and the anode;
d) the potential of the copper.

7.5

The two electrodes (I) and (II) are immersed in neutral solution and separated by a frit. The oxygen partial pressure in the compartment of electrode (II) is equal to $P_{O2,II}$ = 1 bar; that for the other electrode is $P_{O2,I}$ = 0.01 bar. Calculate the open-circuit potential difference between the two electrodes for the following specific cases:
a) the two electrodes are made of noble metal and are both reversible with respect to oxygen;
b) the two electrodes are made of brass and can corrode, with the corrosion rate being controlled by oxygen mass transport, and with the oxygen being reduced at the limiting current. The anodic dissolution of brass obeys Tafel's law, with a Tafel coefficient of β_c = 17 mV.

7.6

The organic coating of a steel water pipe, buried in clay soil, has a circular defect with a diameter of 1 cm. The corrosion potential at the defect is equal to $E_{cor,I}$ = +0.1 V.

© 2007, first edition, EPFL Press

The wall of the pipe has a thickness of 6 mm. The soil resistance is equal to 3000 Ω cm. Because of a construction error, the pipe is in contact with the reinforcement of a building's concrete foundation, thus forming a galvanic cell. The corrosion potential of the metal reinforcement is equal to $E_{cor,II} = +0.6$ V. How long will it take the corrosion to pierce a hole in the pipe?
In order to solve this problem, we assume that the corrosion takes place under ohmic control. The internal resistance of the cell is equal to the resistance of the soil in the vicinity of the defect (Section 12.4):

$$R_{\Omega} = \frac{1}{2\,\kappa\,d}$$

In this expression, κ represents the soil conductivity in Ω^{-1} cm^{-1} and d is the diameter of the defect.

7.7

Estimate the average chromium concentration (weight percent) in a stainless steel of type 18-8 with 0.4% carbon after the sample has undergone a sensitizing heat treatment that resulted in the precipitation of chromium carbide, $Cr_{23}C_6$.

7.8

The propagation rate of a one-dimensional pit in the wall of a steel pipe immersed in an NaCl electrolyte is controlled by the mass transport of ferrous ions. Calculate the time (in hours) that it takes for a hole to pierce the wall. The wall thickness is 3 mm. The solubility of $FeCl_2$ in water is equal to 4.25 mol l^{-1}. For the calculation, we assume an effective diffusion coefficient of ferrous ions of $D_{eff(Fe2+)} = 0.72 \times 10^{-5}$ cm^2 s^{-1}, that includes the influence of migration (cf. Chapter 4).

7.9

The pitting potential of a passive alloy in aerated seawater at pH 8 and at 40 °C is equal to 0.80 V relative to the saturated calomel electrode. Is there a risk of pitting corrosion if this alloy is exposed to aerated seawater at this temperature?

Chapter 8

8.1

Low temperature oxidation of a metal exposed to a low partial pressure of oxygen is studied by the method of photo-electron spectroscopy. The following results are found for the variation of the thickness L as a function of the exposure time t:

© 2007, first edition, EPFL Press

t (min)	10	11	98	1050	9960
L (nm)	1.5	2.0	2.6	3.1	3.6

Determine the growth law and estimate the oxide thickness after 10 years of exposure.

8.2

The relative humidity of an atmosphere at 25 °C is found to be 83%. Calculate the maximum diameter of a pore in which condensation can occur.

8.3

The relative humidity of air at 30°C is found to be 53%. The temperature drops to 20 °C. What is now the relative humidity? Is this greater or smaller than the equilibrium value in contact with a saturated $ZnSO_4$ solution?

8.4

Steel samples with a surface of 25 cm^2 are exposed to the atmosphere and the corroded mass is measured as a function of exposure time by gravimetry. The following results are obtained

Exposure time (months)	Corroded metal mass (mg)
2	20
12	69
22	101
36	148

Estimate the thickness corroded after 20 years; discuss the validity of this extrapolation.

8.5

a) Calculate the total concentration of ferrous ions (Fe^{2+} + $FeOH^+$) in equilibrium with ferrous hydroxide $Fe(OH)_2$ in an electrolyte of pH 6 (T = 25 °C). Which one is the dominating species?
b) In the same fashion, calculate the total concentration of trivalent cations in equilibrium with ferric hydroxide ($Fe(OH)_3$) and identify the dominating species.
c) Anodic dissolution of iron usually produces divalent species. During atmospheric corrosion, the reaction of dissolved divalent species with oxygen, leads to precipitation of $Fe(OH)_3$. Explain this phenomenon.
The following table gives the solubility constants of the different species (W. Stumm, G.F. Lee, *Ind. Engr. Chemistry 53*, 143 (1961):

© 2007, first edition, EPFL Press

Reaction number	Oxidation state of iron	Equation	Equilibrium constant
1	Fe(II)	$Fe(OH)_2 = Fe^{2+} + 2\ OH^-$	$K_1 = 8 \times 10^{-16}$
2	Fe(II)	$Fe(OH)_2 = FeOH^+ + OH^-$	$K_2 = 4 \times 10^{-10}$
3	Fe(III)	$Fe(OH)_3 = Fe^{3+} + 3OH^-$	$K_3 = 10^{-36}$
4	Fe(III)	$Fe(OH)_3 = FeOH^{2+} + 2OH^-$	$K_4 = 6.8 \times 10^{-25}$
5	Fe(III)	$Fe(OH)_3 = Fe(OH)_2^+ + OH^-$	$K_5 = 1.7 \times 10^{-15}$

8.6

Estimate the maximum rate of atmospheric corrosion (mm per year) of a steel plate covered by a water film of 1 μm thickness (25 °C). Why is the corrosion rate observed in practice usually lower than the value calculated in this way?

Chapter 9

9.1

The high-temperature oxidation of nickel obeys the parabolic law. The oxidation constant varies with the oxygen partial pressure according to the relation:

$$k_p = k'_p P_{O2}^{1/n}$$

where k'_p is the rate constant at P_{O2} = 1 bar.
a) Using the results given in Figure 9.13 estimate the activation energy at P_{O2}= 1 bar.

b) Would the activation energy for $P_{O2} = 10^{-2}$ bar be different, and if so, why ?

9.2

Show that, in a monovalent *n*-type semiconducting oxide, the concentration of interstitial cations is proportional to $P_{O2}^{-1/8}$, and that, in a *p*-type monovalent oxide, it is proportional to $P_{O2}^{1/8}$.

9.3

At high temperature, copper oxidizes according to the following reaction:

$$2\ Cu + 1/2\ O_2 \rightarrow Cu_2O$$

Should the presence of a small amount of nickel ions in the oxide increase or decrease the rate of oxidation?

© 2007, first edition, EPFL Press

9.4

Using the thermodynamic data given in Table 2.1, plot $\ln P_{SO2}$ vs. $\ln P_{O2}$ for nickel at 600 °C.
a) Show the regions of stability of Ni, NiO, and NiS.
b) Which of the three solid species is thermodynamically stable in an exhaust gas that contains 2% of SO_2 by volume. The partial pressure of O_2 is 10^{-16} bar, the total pressure is 1 bar.

9.5

A ferritic stainless steel containing (by weight) 23.4% Cr, 0.05% Ti, 1.5% Mo and 0.05% C is annealed at 900 °C. Indicate the concentration of the different carbides (in weight %) that one will find after a rapid quenching of the steel.

Chapter 10

10.1

In a sliding-wear test with a back-and-forth movement, a steel pin rubs against a plate of the same material. The normal force is equal to 100 N and the frequency is 50 Hz. The amplitude of the oscillation is 2 mm. The geometric area of the contact is 3 mm^2, and the friction coefficient is $f = 0.1$.
a) Calculate the thermal energy dissipated per second by the contact.
b) Estimate the increase of the average contact temperature if the thermal conduction length is 1 cm and assuming that the pin and the plate contribute equally to the dissipation of the heat. The thermal conductivity of steel is equal to 0.46 J cm^{-1} s^{-1} K^{-1}.

10.2

According to the data given in Figure 10.15, there is a linear relationship between the hardness of pure metals and their resistance to abrasive wear. Should one expect to find a similar relation for adhesive wear?

10.3

The rate of wear of a non-lubricated steel on steel contact is measured using a pin-on-disk apparatus. The geometric contact area is 4 mm^2. A normal force of 50 N is applied and the rotation of the disk is set to 3 rpm. The pin is located at a distance of 8 cm from the center of the disk. The hardness of the steel is 250 kg mm^{-2}, its thermal conductivity is 0.45 J cm^{-1} K^{-1} s^{-1}, and its heat capacity is 490 J kg^{-1} K^{-1}. Locate the conditions of this test on the wear map of Figure 10.21.

10.4

Water flows in smooth-walled pipe of 10 meter length and 0.1 m diameter under turbulent flow conditions at a flow rate of 90 l/s. The fluid pressures at the entry and

© 2007, first edition, EPFL Press

exit are 1.15 bar and 1.0 bar, respectively. The cinematic viscosity of water is 0.01 cm^2/s. Calculate the friction coefficient.

10.5

The critical shear stress that causes erosion corrosion of copper in a pipe containing sea water under turbulent flow conditions is taken to be 9.6 N/m^2. What would be the critical flow velocity in pipes of 5 cm and 20 cm respectively? We assume that the friction coefficient is given by the Blasius relation:

$$f = 0.32\, Re^{-1/4}$$

For this calculation, use the physical properties of pure water: $\rho = 10^{-3}$ kg m^{-3}; $\nu = 10^{-6}$ m^2 s^{-1}.

10.6

Due to surface rugosity, two steel surfaces sliding on each other in a non-lubricated pin-on disk contact only touch at the asperities. It has been proposed (ref. [10], Chapter 10) that the instantaneous temperature at the asperities T_{inst} can be estimated from the approximate formula:

$$T_{inst} - T_s = \frac{f\beta}{2N^{1/2}} F^{*1/2}\, v^*$$

Here, β is a constant whose value is $\beta \approx 1$, f is the friction coefficient and v^* is the dimensionless sliding velocity. The steady-state temperature of the surface, T_s, is given by equation (10.11). The number of contact points per given geometric surface, N, depends on the non-dimensional pressure F^* according to the relation:

$$N \approx 1 + (4 \times 10^{-3})\, F^*(1 - F^*)$$

Plot the isotherm $T_{inst} - T_s = 400$ °C in a log F^*-log v^* diagram assuming a friction coefficient of $f = 0.5$. Discuss your result.

Chapter 11

11.1

A copper wire is cold worked, resulting in a stored deformation energy of 12 J g^{-1}. An electrochemical cell is established between this copper wire and an annealed copper sheet. Calculate the equilibrium potential of the cell. Is the cold-worked metal at a positive or negative potential relative to that of the annealed electrode? Which one of the two metals will corrode preferentially?

11.2

The minimum stress intensity that can lead to crack growth in a steel used in the construction of a steam turbine rotor corresponds to $K_{ISCC} = 5$ M N $m^{-3/2}$. The

© 2007, first edition, EPFL Press

material is subjected to a tensile stress of 700 MPa. Calculate the maximum admissible length of a crack. We assume that $f_{(a/w)} = 1$.

11.3

A bar of 5 mm diameter made of high strength steel (yield strength $\sigma_e = 1100$ MPa) is subjected to a tension force of 1.5×10^4 N. Indicate if the bar will withstand this force (a) in a non-corrosive environment and (b) in the presence of dissolved hydrogen resulting from a bad zinc coating (cf. Fig. 11.27).

11.4

Using the data from Table 11.20, estimate the solubility of hydrogen in α-iron at 200 °C. The pressure of the hydrogen gas is 20 bar. Give your answer in units of cm^3 H_2 (25 °C, 1 bar) per cm^3 Fe.

11.5

A two-compartment electrochemical cell is used to measure the diffusion coefficient of hydrogen in iron. The solubility of hydrogen in the metal under the conditions of the experiment is equal to 1.62 cm^3 (25 °C, 1 bar) per gram of Fe. At the anode, a current density of 80 μA cm^{-2} is measured, which is attributed to the oxidation of hydrogen. The thickness of the iron sheet is 20 μm. Calculate the diffusion coefficient of hydrogen in iron.

Chapter 12

12.1

The electroless deposition of nickel is studied in a phosphorous acid electrolyte. The process takes place according to the following reaction:

$$Ni^{2+} + H_2PO_2^- + H_2O \rightarrow Ni + H_2PO_3^- + 2\ H^+$$

The anodic partial reaction (hypophosphite) and the cathodic partial reaction (nickel) obey Tafel kinetics:

anodic: $E = 0.085 + 0.065 \ln i_{a,HP}$
cathodic: $E = -0.542 - 0.039 \ln |i_{c,Ni}|$

Here the subscript HP refers to hypophosphite. Calculate the time required to form a 12 μm thick deposit.

12.2

The chemical analysis of drinking water reveals the following cationic concentrations: Ca^{2+}: 10^{-4} mol l^{-1}; Mg^{2+}: 4×10^{-5} mol l^{-1}; Na^+: 10^{-3} mol l^{-1}; NH_4^+: 10^{-5} mol l^{-1}.

© 2007, first edition, EPFL Press

Indicate the hardness of the water in degrees on the French hardness scale (equivalent of 10 mg $CaCO_3$ = 1°F) .

12.3

The adsorption of BTA on iron in sulfuric acid obeys the Langmuir isotherm. The Gibbs free energy of adsorption is given by the following expression (b_L is the adsorption coefficient):

$$\Delta G^{\circ}_{ads} = -RT \ln b_L$$

Use the results from Figure 12.29 to estimate the value of ΔG°_{ads}.
Show that the value of the adsorption coefficient decreases with temperature if the adsorption is exothermic.

12.4

A steel propeller of a boat turns at the speed of 50 rpm in seawater. It undergoes uniform corrosion at the rate of 10 mg day^{-1} cm^{-2}. The surface of the propeller is 2 m^2. Estimate the current required to cathodically protect it.

12.5

Calculate the protection potential, relative to the copper sulfate electrode, of a steel pipe buried in a soil of pH 8. The temperature is 15 °C.

12.6

A steel tank having a surface of 9 m^2 is cathodically protected with a magnesium anode. The polarization resistance of the tank is equal to 4×10^4 Ω cm^2, the anodic Tafel coefficient is β_a = 20 mV. The cathodic partial reaction is the reduction of oxygen at the limiting current. The protection must last 6 years. Because of different losses, only 50% of the anode mass can be used for the protection of the tank. Calculate the mass of the anode that is required.

12.7

A 30 cm diameter straight pipe is buried in soil at a depth of 2 meters. The resistivity of the soil is equal to ρ_e = 3000 Ω cm. A stray current enters the pipe. Its magnitude is determined by placing two copper sulfate reference electrodes at the surface of the soil, one at a position directly perpendicular to the pipe, and the other at a distance of 20 m to the side. A potential difference of 1.25 volts is measured between the two electrodes, with the electrode placed above the pipe as the negative pole. Show that the stray current that enters the pipe is given by:

© 2007, first edition, EPFL Press

$$I = \frac{2\pi L \Delta\Phi}{\rho_e \ln \frac{h^2 + y^2}{h^2}}$$

In this expression, L is the length of the pipe, $\Delta\Phi$ is the potential difference measured between the two reference electrodes, y is the distance between electrodes, and h is the depth to which the pipe is buried. Calculate the stray current density for the given conditions.

12.8

The effect of an inhibitor on the corrosion rate of iron in an acidic environment of pH 0.2 is studied. In the absence of the inhibitor, the anodic and cathodic partial reactions obey the following equations (logarithms are base-10):

anodic: $E = 0.08 + 0.05 \log i_{a,Fe}$
cathodic: $E = -0.36 - 0.12 \log |i_{c,H}|$

a) Calculate the corrosion current density and the corrosion potential.
b) Calculate the exchange current density of the cathodic partial reaction (proton reduction).
c) The presence of an inhibitor decreases the exchange current density of the partial cathodic reaction by a factor of 1000. Calculate the corrosion current density and the corrosion potential.
d) Calculate the inhibition efficiency.

© 2007, first edition, EPFL Press

A.4 LIST OF SYMBOLS

A	surface (m^2)
a_i	activity of species i
B_i	chemical species i
b_L	adsorption coefficient (Langmuir isotherm)
b_a, (b_c)	anodic (cathodic) Tafel coefficients, base 10 logarithm (V)
C	capacitance ($C\ V^{-1}$ or $C\ V^{-1}\ m^{-2}$)
c_i	concentration of the species i ($mol\ m^{-3}$)
c_{no}	electron concentration in an intrinsic semiconductor ($mol\ m^{-3}$)
c_{po}	hole concentration in an intrinsic semiconductor ($mol\ m^{-3}$)
D	diffusion coefficient ($m^2\ s^{-1}$)
D_h	hydraulic diameter (m)
E	electrode potential relative to the standard hydrogen electrode (V)
$E_{(sce)}$	electrode potential relative to the saturated calomel electrode (V)
E_{rev}	reversible potential (V)
$E°$	standard potential (V)
E_{cor}	corrosion potential (V)
E_{prot}	protection potential (V)
E_{kin}	kinetic energy (J)
E_F	Fermi energy (eV)
E_g	band gap energy (eV)
E_b	binding energy
F	Faraday constant ($C\ mol^{-1}$)
f	activity coefficient
f	friction coefficient (dimensionless)
g	gravitational constant ($m\ s^{-2}$)
h	Planck's constant (J s)
h	height (m)
H	enthalpy ($J\ mol^{-1}$)
i	current density ($A\ m^{-2}$)
i_0	exchange current density ($A\ m^{-2}$)
i_{cor}	corrosion current density ($A\ m^{-2}$)
i_l	limiting current density ($A\ m^{-2}$)

© 2007, first edition, EPFL Press

i_a (i_c)	anodic (cathodic) current density (A m^{-2})
$i_{a,k}$ ($i_{c,k}$)	anodic (cathodic) current density of a partial reaction k (A m^{-2})
J	ionic strength (mol m^{-3})
K	equilibrium constant
K_I	mode I stress intensity factor (MPa m$^{-1/2}$)
K_{IC}	fracture toughness (MPa m$^{-1/2}$)
K_{ISCC}	limiting stress intensity factor for stress corrosion cracking (MPa m$^{-1/2}$)
k	rate constant
k_L	mass transport coefficient (mol m^{-1} s^{-1})
L	characteristic length (m)
L	thickness (m)
m	mass (kg)
M	molar mass (kg mol^{-1})
n	charge number (dimensionless)
n_i	amount of substance i (mol)
N	number of cycles
N_f	number of cycles before failure
N_L	Avogadro's number
N_0	number of adsorption sites (m^{-2})
N_i	number of adsorption sites occupied by species i (m^{-2})
N_i	flux of the species i (mol m^{-2} s^{-1})
P	pressure (N m^{-2})
P_i	partial pressure of species i (N m^{-2})
P_i^r	permeability (mol m^{-1} s^{-1})
q	charge density (C m^{-2})
Q	electric charge (C)
Q_{ads}	heat of adsorption (J mol^{-1})
r	radius (m)
r_p	polarization resistance $r_p = dE/di$ (Ω m^2)
R	gas constant (J mol^{-1} K^{-1})
Re	Reynolds number (dimensionless)
R_Ω	ohmic resistance (Ω)
R_a	average rugosity (μm)
s_K	condensation coefficient (dimensionless)
$S°$	entropy of formation (J mol^{-1} K^{-1})
Sh	Sherwood number
Sc	Schmidt number
t	time (s)
T	temperature (K)
t_i	transport number of an ion (dimensionless)
U	cell voltage (V)
u_i	mobility of ion i (m mol N^{-1} s^{-1})
v	flow velocity (m s^{-1})

© 2007, first edition, EPFL Press

v	reaction rate (mol m^{-2} s^{-1})
v_{cor}	corrosion rate (mol m^{-2} s^{-1})
V	volume (m^3)
V^m	molar volume
X_i	mole fraction of species i (dimensionless)
β_a (β_c)	anodic (cathodic) Tafel coefficient (V)
γ	surface tension (N m^{-1}) or surface energy (J m^{-2})
δ	thickness of the diffusion layer (m)
ΔG	free energy of reaction (J mol^{-1})
$\Delta G°$	standard free energy of reaction (J mol^{-1})
$\Delta G^{\#}$	free energy of activation (J mol^{-1})
ΔH	enthalpy of reaction (J mol^{-1})
$\Delta H°$	standard enthalpy of reaction (J mol^{-1})
ΔS	entropy of reaction (J mol^{-1} K^{-1})
$\Delta S°$	standard entropy of reaction (J mol^{-1} K^{-1})
$\Delta\Phi_{JC}$	liquid junction potential (V)
$\Delta\Phi_{\Omega}$	ohmic potential difference (V)
ε	dielectric constant (dimensionless number)
ε_0	permittivity constant (C^2 N^{-1} m^{-2})
ξ	polarization: $\xi = E - E_{cor}$ (V)
η	overpotential: $\eta = E - E_{rev}$ (V)
Θ	contact angle
θ_i	surface coverage of species i (dimensionless)
κ	conductivity (Ω^{-1} m^{-1})
λ	electron mean free path (m)
λ	potential sweep rate (V s^{-1})
Λ	electron depth (m)
μ_i	chemical potential of species i (J mol^{-1})
μ_i°	standard chemical potential of species i (J mol^{-1})
μ_i^*	electrochemical potential of species i : $\mu_i^* = \mu_i + z_i F\Phi$ (J mol^{-1})
ν	frequency (s^{-1})
ν	kinematic viscosity (m^2 s^{-1})
ν_i	stoichiometric coefficient of species i (dimensionless)
ρ	density (kg m^{-3})
ρ_e	resistivity (Ω m)
Γ_i	surface excess of species i (mol m^{-2})
Φ	electric potential, or Galvani potential (V)
Φ_σ	Galvani potential in the electrolyte (V)
Φ_m	Galvani potential in the metal (V)
Φ_w	work function (J)
ω	angular velocity (rad s^{-1})

© 2007, first edition, EPFL Press

Subscripts

a	anodic
ads	adsorption, adsorbed
aq	hydrated
B	chemical species B
b	bulk
c	cathodic
cb	conduction band
ch	chemical
cor	corrosion
des	desorption
e	electron
gr	growth
GC	Gouy-Chapman
H	Helmholtz
i	chemical species *i*
k	partial reaction *k*
m	metal
ox	species "oxidized" in a half-cell reaction: $\Sigma \nu_{i,ox} B_{i,ox} + ne = \Sigma \nu_{i,red} B_{i,red}$
oxd	oxide
red	species "reduced" in a half-cell reaction: $\Sigma \nu_{i,ox} B_{i,ox} + ne = \Sigma \nu_{i,red} B_{i,red}$
s	surface
sat	saturated, saturation
sc	space charge
sl	sliding
rev	reversible (at equilibrium)
vb	valence band
Ω	ohmic
+ (–)	cation or anion in a binary electrolyte
(g)	gas phase
(l)	liquid phase
(s)	solid phase

Superscripts

°	standard state
m	partial molar quantity

© 2007, first edition, EPFL Press