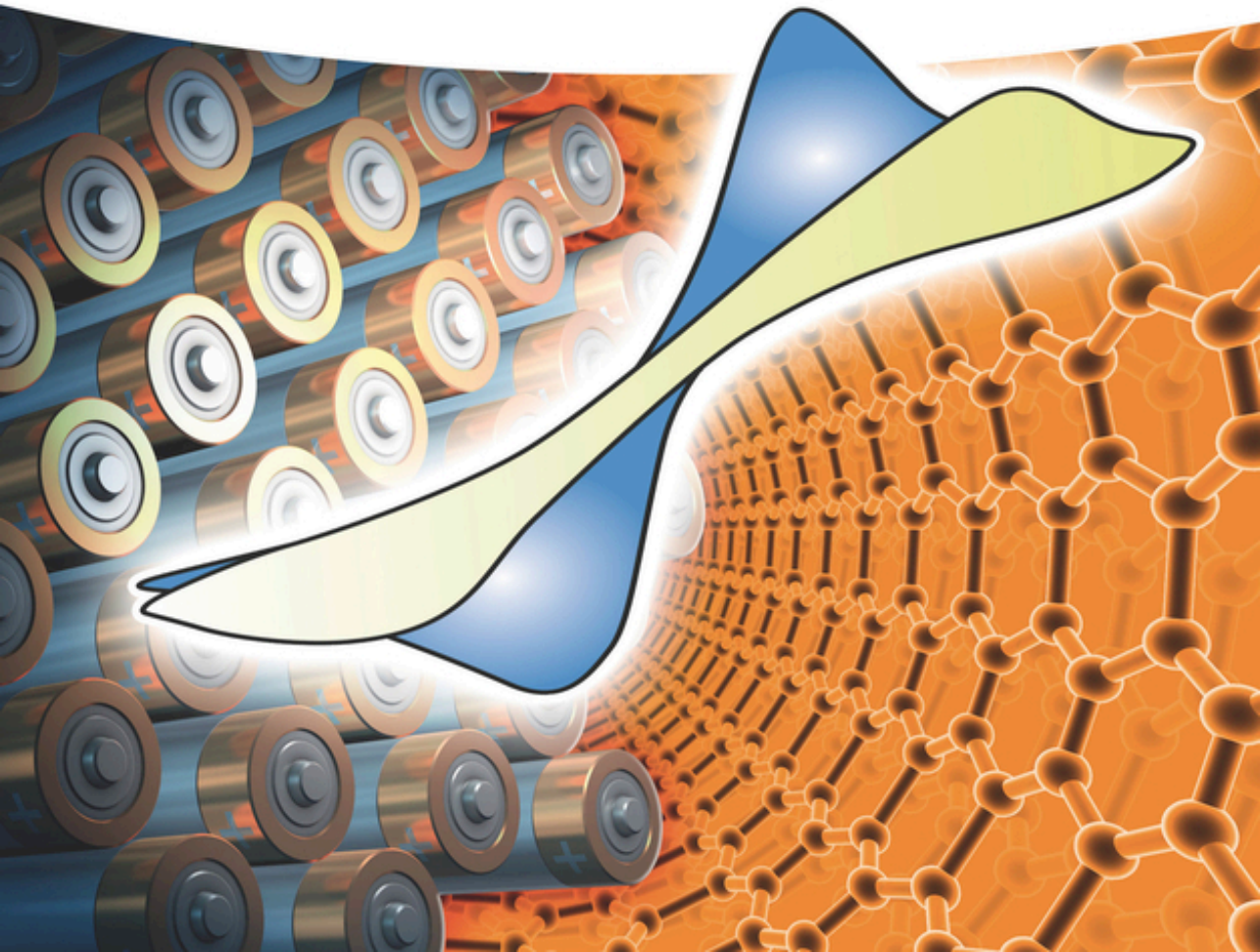
WILEY-VCH

Noam Eliaz and Eliezer Gileadi

Physical Electrochemistry

Fundamentals, Techniques, and Applications

Second, Completely Revised and Updated Edition



Physical Electrochemistry

Physical Electrochemistry

Fundamentals, Techniques, and Applications

Noam Eliaz and Eliezer Gileadi

Second Edition

WILEY-VCH

Authors
Prof. Noam Eliaz
Tel Aviv University
Department of Materials Science & Engineering
Room 121, Wolfson Building
6997801 Tel Aviv
Israel

Prof. Dr. Eliezer Gileadi
Tel Aviv University
School of Chemistry
Shenkarbuilding
6997801 Tel Aviv
Israel

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>.

Print ISBN: 978-3-527-34139-9
ePDF ISBN: 978-3-527-34140-5
ePub ISBN: 978-3-527-34142-9
oBook ISBN: 978-3-527-34143-6

Cover Design Formgeber, Mannheim, Germany
Typesetting SPi Global, Chennai, India
Printing and Binding

Printed on acid-free paper

10 9 8 7 6 5 4 3 2 1

Dedicated to our beloved wives, Dalia Papouchado and Billie Eliaz, for their love, continued support, encouragement, and patience. Without them, our scientific careers would not be what they have been, and this book would not have been written.

Contents

Preface *xvii*
Symbols and Abbreviations *xix*

1 Introduction *1*
1.1 General Considerations *1*
1.1.1 The Transition from Electronic to Ionic Conduction *1*
1.1.2 The Resistance of the Interface can be Infinite *2*
1.1.3 Mass-Transport Limitation *2*
1.1.4 The Capacitance at the Metal/Solution Interphase *4*
1.2 Polarizable and Nonpolarizable Interfaces *4*
1.2.1 Phenomenology *4*
1.2.2 The Equivalent Circuit Representation *5*
Further Reading *7*

2 The Potentials of Phases *9*
2.1 The Driving Force *9*
2.1.1 Definition of the Electrochemical Potential *9*
2.1.2 Separability of the Chemical and the Electrical Terms *10*
2.2 Two Cases of Special Interest *11*
2.2.1 Equilibrium of a Species Between two Phases in Contact *11*
2.2.2 Two Identical Phases not at Equilibrium *12*
2.3 The Meaning of the Standard Hydrogen Electrode (SHE) Scale *13*
Further Reading *15*

3 Fundamental Measurements in Electrochemistry *17*
3.1 Measurement of Current and Potential *17*
3.1.1 The Cell Voltage is the Sum of Several Potential Differences *17*
3.1.2 Use of a Nonpolarizable Counter Electrode *17*
3.1.3 The Three-Electrode Setup *18*
3.1.4 Residual jR_S Potential Drop in a Three-Electrode Cell *18*
3.2 Cell Geometry and the Choice of the Reference Electrode *19*
3.2.1 Types of Reference Electrodes *19*
3.2.2 Use of an Auxiliary Reference Electrode for the Study of Fast Transients *20*

3.2.3 Calculating the Uncompensated Solution Resistance for a few Simple Geometries *21*
3.2.3.1 Planar Configuration *21*
3.2.3.2 Cylindrical Configuration *21*
3.2.3.3 Spherical Symmetry *22*
3.2.4 Positioning the Reference Electrode *22*
3.2.5 Edge Effects *24*
Further Reading *26*

4 Electrode Kinetics: Some Basic Concepts *27*
4.1 Relating Electrode Kinetics to Chemical Kinetics *27*
4.1.1 The Relation of Current Density to Reaction Rate *27*
4.1.2 The Relation of Potential to Energy of Activation *28*
4.1.3 Mass-Transport Limitation Versus Charge-Transfer Limitation *30*
4.1.4 The Thickness of the Nernst Diffusion Layer *31*
4.2 Methods of Measurement *33*
4.2.1 Potential Control Versus Current Control *33*
4.2.2 The Need to Measure Fast Transients *35*
4.2.3 Polarography and the Dropping Mercury Electrode (DME) *37*
4.3 Rotating Electrodes *40*
4.3.1 The Rotating Disk Electrode (RDE) *40*
4.3.2 The Rotating Cone Electrode (RConeE) *44*
4.3.3 The Rotating Ring Disk Electrode (RRDE) *45*
Further Reading *47*

5 Single-Step Electrode Reactions *49*
5.1 The Overpotential, η *49*
5.1.1 Definition and Physical Meaning of Overpotential *49*
5.1.2 Types of Overpotential *51*
5.2 Fundamental Equations of Electrode Kinetics *52*
5.2.1 The Empirical Tafel Equation *52*
5.2.2 The Transition-State Theory *53*
5.2.3 The Equation for a Single-Step Electrode Reaction *54*
5.2.4 Limiting Cases of the General Equation *56*
5.3 The Symmetry Factor, β, in Electrode Kinetics *59*
5.3.1 The Definition of β *59*
5.3.2 The Numerical Value of β *60*
5.4 The Marcus Theory of Charge Transfer *61*
5.4.1 Outer-Sphere Electron Transfer *61*
5.4.2 The Born–Oppenheimer Approximation *62*
5.4.3 The Calculated Energy of Activation *63*
5.4.4 The Value of β and its Potential Dependence *64*
5.5 Inner-Sphere Charge Transfer *65*
5.5.1 Metal Deposition *65*
Further Reading *66*

6 Multistep Electrode Reactions *67*
6.1 Mechanistic Criteria *67*
6.1.1 The Transfer Coefficient, α, and its Relation to the Symmetry Factor, β *67*
6.1.2 Steady State and Quasi-Equilibrium *69*
6.1.3 Calculation of the Tafel Slope *71*
6.1.4 Reaction Orders in Electrode Kinetics *74*
6.1.5 The Effect of pH on Reaction Rates *77*
6.1.6 The Enthalpy of Activation *79*
Further Reading *81*

7 Specific Examples of Multistep Electrode Reactions *83*
7.1 Experimental Considerations *83*
7.1.1 Multiple Processes in Parallel *83*
7.1.2 The Level of Impurity that can be Tolerated *84*
7.2 The Hydrogen Evolution Reaction (HER) *87*
7.2.1 Hydrogen Evolution on Mercury *87*
7.2.2 Hydrogen Evolution on Platinum *89*
7.3 Possible Paths for the Oxygen Evolution Reaction *91*
7.4 The Role and Stability of Adsorbed Intermediates *94*
7.5 Adsorption Energy and Catalytic Activity *95*
Further Reading *96*

8 The Electrical Double Layer (EDL) *97*
8.1 Models of Structure of the EDL *97*
8.1.1 Phenomenology *97*
8.1.2 The Parallel-Plate Model of Helmholtz *99*
8.1.3 The Diffuse Double Layer Model of Gouy and Chapman *100*
8.1.4 The Stern Model *103*
8.1.5 The Role of the Solvent at the Interphase *105*
Further Reading *107*

9 Electrocapillary *109*
9.1 Thermodynamics *109*
9.1.1 Adsorption and Surface Excess *109*
9.1.2 The Gibbs Adsorption Isotherm *111*
9.1.3 The Electrocapillary Equation *112*
9.2 Methods of Measurement and Some Results *114*
9.2.1 The Electrocapillary Electrometer *114*
9.2.2 Some Experimental Results *119*
9.2.2.1 The Adsorption of Ions *119*
9.2.2.2 Adsorption of Neutral Molecules *120*
Further Reading *122*

10 Intermediates in Electrode Reactions *123*
10.1 Adsorption Isotherms for Intermediates Formed by Charge Transfer *123*

10.1.1 General *123*
10.1.2 The Langmuir Isotherm and its Limitations *123*
10.1.3 Application of the Langmuir Isotherm for Charge-Transfer Processes *125*
10.1.4 The Frumkin Adsorption Isotherms *126*
10.2 The Adsorption Pseudocapacitance C_ϕ *127*
10.2.1 Formal Definition of C_ϕ and its Physical Understanding *127*
10.2.2 The Equivalent-Circuit Representation *129*
10.2.3 Calculation of C_ϕ as a function of θ and E *130*
Further Reading *133*

11 Underpotential Deposition and Single-Crystal Electrochemistry *135*
11.1 Underpotential Deposition (UPD) *135*
11.1.1 Definition and Phenomenology *135*
11.1.2 UPD on Single Crystals *139*
11.1.3 Underpotential Deposition of Atomic Oxygen and Hydrogen *141*
Further Reading *142*

12 Electrosorption *145*
12.1 Phenomenology *145*
12.1.1 What is Electrosorption? *145*
12.1.2 Electrosorption of Neutral Organic Molecules *147*
12.1.3 The Potential of Zero Charge, E_{pzc}, and its Importance in Electrosorption *148*
12.1.4 The *Work Function* and the Potential of Zero Charge *151*
12.2 Adsorption Isotherms for Neutral Species *152*
12.2.1 General Comments *152*
12.2.2 The Parallel-Plate Model of Frumkin et al. *153*
12.2.3 The Water Replacement Model of Bockris et al. *155*
Further Reading *157*

13 Fast Transients, the Time-Dependent Diffusion Equation, and Microelectrodes *159*
13.1 The Need for Fast Transients *159*
13.1.1 General *159*
13.1.2 Small-Amplitude Transients *161*
13.1.3 The Sluggish Response of the Electrochemical Interphase *162*
13.1.4 How can the Slow Response of the Interphase be Overcome? *162*
13.1.4.1 Galvanostatic Transients *162*
13.1.4.2 The Double-Pulse Galvanostatic Method *163*
13.1.4.3 The Coulostatic (Charge-Injection) Method *164*
13.2 The Diffusion Equation *167*
13.2.1 The Boundary Conditions of the Diffusion Equation *167*
13.2.1.1 Potential Step, Reversible Case (Chrono-Amperometry) *168*
13.2.1.2 Potential Step, High Overpotential Region (Chrono-Amperometry) *171*

13.2.1.3 Current Step (Chronopotentiometry) *172*
13.3 Microelectrodes *174*
13.3.1 The Unique Features of Microelectrodes *174*
13.3.2 Enhancement of Diffusion at a Microelectrode *175*
13.3.3 Reduction of the Solution Resistance *176*
13.3.4 The Choice between Single Microelectrodes and Large Ensembles *176*
Further Reading *178*

14 Linear Potential Sweep and Cyclic Voltammetry *181*
14.1 Three Types of Linear Potential Sweep *181*
14.1.1 Very Slow Sweeps *181*
14.1.2 Studies of Oxidation or Reduction of Species in the Bulk of the Solution *182*
14.1.3 Studies of Oxidation or Reduction of Species Adsorbed on the Surface *182*
14.1.4 Double-Layer Charging Currents *183*
14.1.5 The Form of the Current–Potential Relationship *185*
14.2 Solution of the Diffusion Equations *186*
14.2.1 The Reversible Region *186*
14.2.2 The High-Overpotential Region *187*
14.3 Uses and Limitations of the Linear Potential Sweep Method *188*
14.4 Cyclic Voltammetry for Monolayer Adsorption *190*
14.4.1 Reversible Region *190*
14.4.2 The High-Overpotential Region *192*
Further Reading *193*

15 Electrochemical Impedance Spectroscopy (EIS) *195*
15.1 Introduction *195*
15.2 Graphical Representations *200*
15.3 The Effect of Diffusion Limitation – The Warburg Impedance *203*
15.4 Advantages, Disadvantages, and Applications of EIS *206*
Further Reading *211*

16 The Electrochemical Quartz Crystal Microbalance (EQCM) *213*
16.1 Fundamental Properties of the EQCM *213*
16.1.1 Introduction *213*
16.1.2 The EQCM *214*
16.1.3 The Effect of Viscosity *217*
16.1.4 Immersion in a Liquid *218*
16.1.5 Scales of Roughness *218*
16.2 Impedance Analysis of the EQCM *219*
16.2.1 The Extended Equation for the Frequency Shift *219*
16.2.2 Other Factors Influencing the Frequency Shift *220*

16.3 Uses of the EQCM as a Microsensor *220*
16.3.1 Advantages and Limitations *220*
16.3.2 Some Applications of the EQCM *222*
Further Reading *225*

17 Corrosion *227*
17.1 The Definition of Corrosion *227*
17.2 Corrosion Costs *230*
17.3 Thermodynamics of Corrosion *232*
17.3.1 Introduction and Important Terms *232*
17.3.2 Electrode Potentials and the Standard Electromotive Force (EMF) Series *236*
17.3.3 The Dependence of Free Energy on the Equilibrium Constant and Cell Potential *241*
17.3.4 The Nernst Equation *241*
17.3.5 The Potential–pH (Pourbaix) Diagrams *242*
17.4 Kinetics of Corrosion *252*
17.4.1 Introduction and Important Terms *252*
17.4.2 Two Limiting Cases of the Butler–Volmer Equation: Tafel Extrapolation and Polarization Resistance *255*
17.4.3 Corrosion Rate *257*
17.4.4 The Mixed-Potential Theory and the Evans Diagrams *257*
17.4.5 Passivation and its Breakdown *264*
17.5 Corrosion Measurements *270*
17.5.1 Non-Electrochemical Tests *270*
17.5.2 Electrochemical Tests *272*
17.5.2.1 Open-Circuit Potential (OCP) Measurements *272*
17.5.2.2 Polarization Tests *273*
17.5.2.3 Linear Polarization Resistance (LPR) *277*
17.5.2.4 Zero-Resistance Ammetry (ZRA) *277*
17.5.2.5 Electrochemical Noise (EN) Measurements *278*
17.5.2.6 Electrochemical Hydrogen Permeation Tests *279*
17.5.3 Complementary Surface-Sensitive Analytical Characterization Techniques *284*
17.6 Forms of Corrosion *286*
17.6.1 Uniform (General) Corrosion *286*
17.6.2 Localized Corrosion *289*
17.6.2.1 Crevice Corrosion *289*
17.6.2.2 Filiform Corrosion *291*
17.6.2.3 Pitting Corrosion *291*
17.6.3 Intergranular Corrosion *293*
17.6.3.1 Sensitization *293*
17.6.3.2 Exfoliation *294*
17.6.4 Dealloying *295*
17.6.5 Galvanic (Bimetallic) Corrosion *295*
17.6.6 Environmentally Induced Cracking (EIC)/Environment-Assisted Cracking (EAC) *297*

17.6.6.1 Hydrogen Embrittlement (HE) *297*
17.6.6.2 Hydrogen-Induced Blistering *299*
17.6.6.3 Hydrogen Attack *299*
17.6.6.4 Stress Corrosion Cracking (SCC) *300*
17.6.6.5 Corrosion Fatigue (CF) *303*
17.6.7 Erosion Corrosion *304*
17.6.8 Microbiological Corrosion (MIC) *305*
17.7 Corrosion Protection *308*
17.7.1 Cathodic Protection *308*
17.7.1.1 Cathodic Protection with Sacrificial Anodes *308*
17.7.1.2 Impressed-Current Cathodic Protection (ICCP) *310*
17.7.2 Anodic Protection *312*
17.7.3 Corrosion Inhibitors *313*
17.7.4 Coatings *315*
17.7.5 Other Mitigation Practices *320*
Further Reading *321*

18 Electrochemical Deposition *323*
18.1 Electroplating *323*
18.1.1 Introduction *323*
18.1.2 The Fundamental Equations of Electroplating *324*
18.1.3 Practical Aspects of Metal Deposition *325*
18.1.4 Hydrogen Evolution as a Side Reaction *326*
18.1.5 Plating of Noble Metals *327*
18.1.6 Current Distribution in Electroplating *328*
18.1.6.1 Uniformity of Current Distribution *328*
18.1.6.2 The Faradaic Resistance (R_F) and the Solution Resistance (R_S) *328*
18.1.6.3 The Dimensionless Wagner Number *329*
18.1.6.4 Kinetically Limited Current Density *333*
18.1.7 Throwing Power *334*
18.1.7.1 Macro Throwing Power *334*
18.1.7.2 Micro Throwing Power *334*
18.1.8 The Use of Additives *336*
18.1.9 The Microstructure of Electrodeposits and the Evolution of Intrinsic Stresses *339*
18.1.10 Pulse Plating *341*
18.1.11 Plating from Nonaqueous Solutions *343*
18.1.11.1 Statement of the Problem *343*
18.1.11.2 Methods of Plating Al *345*
18.1.12 Electroplating of Alloys *346*
18.1.12.1 General Observations *346*
18.1.12.2 Some Specific Examples *349*
18.1.13 The Mechanism of Charge Transfer in Metal Deposition *351*
18.1.13.1 Metal Ions Crossing the Interphase Carry the Charge Across it *351*
18.2 Electroless Deposition of Metals *352*
18.2.1 Some Fundamental Aspects of Electroless Plating of Metals and Alloys *352*

18.2.2 The Activation Process *353*
18.2.3 The Reducing Agent *353*
18.2.4 The Complexing Agent *354*
18.2.5 The Mechanism of Electroless Deposition *354*
18.2.6 Advantages and Disadvantages of Electroless Plating Compared to Electroplating *357*
18.3 Electrophoretic Deposition (EPD) *358*
Further Reading *361*

19 Electrochemical Nanotechnology *363*
19.1 Introduction *363*
19.2 Nanoparticles and Catalysis *363*
19.2.1 Surfaces and Interfaces *364*
19.2.2 The Vapor Pressure of Small Droplets and the Melting Point of Solid NPs *365*
19.2.3 The Thermodynamic Stability and Thermal Mobility of NPs *368*
19.2.4 Catalysts *368*
19.2.5 The Effect of Particle Size on Catalytic Activity *369*
19.2.6 Nanoparticles Compared to Microelectrodes *370*
19.2.7 The Need for High Surface Area *371*
19.3 Electrochemical Printing *372*
19.3.1 Electrochemical Printing Processes *373*
19.3.2 Nanoelectrochemistry Using Micro- and Nano-Electrodes/Pipettes *379*
Further Reading *384*

20 Energy Conversion and Storage *387*
20.1 Introduction *387*
20.2 Batteries *388*
20.2.1 Classes of Batteries *388*
20.2.2 The Theoretical Limit of Energy per Unit Weight *390*
20.2.3 How is the Quality of a Battery Defined? *391*
20.2.4 Primary Batteries *392*
20.2.4.1 Why Do We Need Primary Batteries? *392*
20.2.4.2 The Leclanché and the Alkaline Batteries *392*
20.2.4.3 The Li–Thionyl Chloride Battery *393*
20.2.4.4 The Lithium–Iodine Solid-State Battery *395*
20.2.5 Secondary Batteries *396*
20.2.5.1 Self-Discharge and Specific Energy *396*
20.2.5.2 Battery Stacks Versus Single Cells *396*
20.2.5.3 Some Common Types of Secondary Batteries *397*
20.2.5.4 The Li-ion Battery *402*
20.2.5.5 Metal–Air Batteries *408*
20.2.6 Batteries-Driven Electric Vehicles *409*
20.2.7 The Polarity of Batteries *410*
20.3 Fuel Cells *412*
20.3.1 The Specific Energy of Fuel Cells *412*

20.3.2 The Phosphoric Acid Fuel Cell (PAFC) *412*
20.3.3 The Direct Methanol Fuel Cell (DMFC) *415*
20.3.4 The Proton Exchange Membrane Fuel Cell (PEMFC) *418*
20.3.5 The Alkaline Fuel Cell (AFC) *420*
20.3.6 High-Temperature Fuel Cells *421*
20.3.6.1 The Solid Oxide Fuel Cell (SOFC) *421*
20.3.6.2 The Molten Carbonate Fuel Cell (MCFC) *422*
20.3.7 Porous Gas Diffusion Electrodes *423*
20.3.8 Fuel-Cell-Driven Vehicles *426*
20.3.9 Criticism of the Fuel Cells Technology *427*
20.4 Supercapacitors *428*
20.4.1 Electrostatic Considerations *428*
20.4.2 The Energy Stored in a Capacitor *429*
20.4.3 The Essence of Supercapacitors *430*
20.4.4 Advantages of Supercapacitors *432*
20.4.5 Barriers for Supercapacitors *435*
20.4.6 Applications of Supercapacitors *435*
20.5 Hydrogen Storage *436*
Further Reading *443*

Index *445*

Preface

Physical electrochemistry deals with the theory of the double layer at the metal/solution interphase, the thermodynamics and kinetics (rates and mechanisms) of reactions and processes that involve electron transfer. It is important in many aspects of fundamental chemistry, physics, biology, and engineering. There are many applications of electrochemistry, including corrosion, electrochemical deposition, electroforming, electromachining and electropolishing, electro-organic synthesis, biosensors, batteries, fuel cells, and supercapacitors.

In spite of its importance, physical electrochemistry is rarely included in the undergraduate curriculum of chemistry and engineering in universities around the world. This book aims to serve as a key textbook in undergraduate courses that deal with electrochemistry, and also as a reference source for graduate students, researchers, and engineers who have interest in the field. Admittedly, the book contains more than what could be taught in one semester. This is deliberate in order to allow some choice for the teacher to concentrate on aspects of the field that best fit the needs of the particular class. However, even covering one half or two thirds of the material in this book should provide students with some understanding of physical electrochemistry, the techniques applied, and at least one of the applications in which electrochemistry is involved, and facilitate the learning and understanding of some specific subject that may be needed later in his or her professional life. This book is also recommended as a text suitable for self-learning, which could be used to introduce scientists and engineers who have not had an opportunity to participate in a formal course on electrochemistry to aspects of this field needed for their research and development.

The book can be divided into three parts: (i) *the fundamentals of electrochemistry*: the potentials of phases, the electrical double layer (EDL), electrode kinetics, single-step and multistep electrode reactions, electrocapillarity, electrosorption, underpotential deposition (UPD), and single-crystal electrochemistry; (ii) the most important *electrochemical measurement techniques*: cyclic voltammetry (CV), rotating-disk electrodes, microelectrodes and nanoelectrodes, electrochemical impedance spectroscopy (EIS), and electrochemical quartz crystal microbalance (EQCM); and (iii) *applications of electrochemistry in materials science and engineering, nanoscience and nanotechnology, and industry*: corrosion, electrochemical deposition (electroplating of metals and alloys, electroless and electrophoretic deposition), nanoparticles and surfaces,

electrocatalysis, electrochemical printing, and energy conversion and storage (batteries, fuel cells, supercapacitors, and hydrogen storage).

The first edition of this book was published by Wiley-VCH in 2011. Following its success, we made our best efforts to revise and improve this book significantly by (i) shortening certain sections that we find less needed for students nowadays; (ii) updating and extending some chapters according to the state-of-the-art in the field, for example, electrochemical printing, batteries, fuels cells, supercapacitors, and hydrogen storage; (iii) adding key illustrations (figures and tables); (iv) adding recommended references at the end of each chapter; etc. Thus, we believe that the second edition will be valuable also to those of you who have read the first edition.

When writing this book we took advantage of our long experience in teaching courses such as physical electrochemistry, corrosion engineering, and materials science and engineering. There are different criteria by which the quality of a textbook could be judged. From our point of view, the success or failure of this book will be judged by its ability to enhance and spread the teaching and use of physical electrochemistry, and establish it as the basis for graduate courses offered widely in universities around the world. We hope you enjoy reading our book and find it easy to follow and enriching!

July 2018

Noam Eliaz and Eliezer Gileadi
Tel Aviv, Israel

Symbols and Abbreviations

Symbols

a	empirical Tafel constant	V
a_i	activity of substance i	mol m^{-3}, mol kg^{-1}, or mole fraction
A	affinity of a reaction	J mol^{-1}
A	exposed surface area	cm^2
A_p	area of the drop during the application of pulse in normal-pulse polarography	m^2
b	Tafel slope	V $decade^{-1}$
b_a	Tafel slope of the anodic (oxidation) reaction	V $decade^{-1}$
b_c	Tafel slope of the cathodic (reduction) reaction	V $decade^{-1}$
B_o	Bond number	dimensionless
c_i	concentration of substance i	mol m^{-3}
c_b	concentration of the electroactive species in the bulk of the solution	mol m^{-3}
c_s	concentration of the electroactive species at the surface ($x = 0$)	mol m^{-3}
$c_{\ddagger}$	concentration of the activated complex	mol m^{-3}
C_{dl}	double-layer capacitance	μF cm^{-2}
C_H	the Helmholtz capacitance	μF cm^{-2}
C_L	the adsorption pseudocapacitance derived from the Langmuir isotherm	μF cm^{-2}
C_0	subsurface concentration of atomic hydrogen	mol m^{-3}
C_{0R}	summation of the subsurface concentration of hydrogen in interstitial lattice sites and reversible trap sites on the charging side of the sample	mol m^{-3}
C_ϕ	adsorption pseudocapacitance	μF cm^{-2}
d	distance between the tip of the Luggin capillary and the working electrode	m
d	distance between the two plates of a capacitor	m
D	diffusion coefficient	m^2 s^{-1}

D_{eff}	effective diffusion coefficient	$\text{m}^2\ \text{s}^{-1}$
D_0	nozzle diameter	m
E	electrical potential	V
E^0	standard potential	V
E_{app}	applied potential	V
E_{b}	breakdown potential	V
E_{corr}	corrosion potential	V
E_{max}	the potential where the pseudocapacitance reaches its maximum value	V
E_{mp}	mixed potential	V
E_{p}	protection potential	V
E_{pp}	primary passivation potential	V
E_{pzc}	potential of zero charge	V
E_{rev}	reversible potential according to the Nernst equation	V
$E_{1/2}$	polarographic half-wave potential	V
f	fugacity	atm
f	Frumkin parameter	dimensionless
f_0	resonance frequency	Hz
G	Gibbs free energy	J mol^{-1}
ΔG	change in the Gibbs free energy	J mol^{-1}
ΔG_{solv}	energy of hydration of a metal ion	J mol^{-1}
ΔG^0	change in the standard Gibbs free energy	J mol^{-1}
$\Delta G^{0\ddagger}$	the standard electrochemical Gibbs free energy of activation	J mol^{-1}
$\Delta \overline{G}^0$	the change in the standard electrochemical Gibbs free energy	J mol^{-1}
ΔG^0_θ	the standard Gibbs free energy of adsorption for a chosen value of θ	J mol^{-1}
I	current	A
I	ionic strength	m or M
I_{d}	diffusion-limited current in dropping mercury electrode	A
I_{D}	current at the disk of a rotating ring disk electrode	A
I_{L}	mass-transport-limited current	A
I_{R}	current at the ring of a rotating ring disk electrode	A
j	current density	A m^{-2}
j_{a}	net anodic current density	A m^{-2}
j_{avg}	average current density in pulse plating	A m^{-2}
j_{c}	net cathodic current density	A m^{-2}
$j_{\text{ac}}/j_{\text{ct}}$	activation-controlled current density/charge-transfer current density	A m^{-2}
j_{cc}	critical current density	A m^{-2}

j_{corr}	corrosion current density	$A\,m^{-2}$
j_{dep}	deposition current density	$A\,m^{-2}$
j_{dl}	double-layer charging current	$A\,m^{-2}$
j_L	limiting current density	$A\,m^{-2}$
$j_{L,chem}$	chemically-controlled limiting current density	$A\,m^{-2}$
j_{max}	the final steady-state current density	$A\,m^{-2}$
j_p	applied peak current density	$A\,m^{-2}$
j_{pas}	passive current density	$A\,m^{-2}$
j_0	exchange current density	$A\,m^{-2}$
J_{ss}	permeation flux of hydrogen at steady state	$mol\,m^{-2}\,s^{-1}$
k	rate constant for a homogenous reaction	depends on order
k_b	homogeneous rate constant for "backward" reaction	depends on order
k_f	homogeneous rate constant for "forward" reaction	depends on order
k^0	chemical (heterogeneous) rate constant at the reversible potential $\Delta\phi = 0$	$m\,s^{-1}$
$k_{s,h}$	heterogeneous rate constant at standard potential	$m\,s^{-1}$
K	equilibrium constant	dimensionless
$K^{\ddagger}$	equilibrium constant for formation of the activated complex	dimensionless
l	characteristic length of rotating disk electrode	m
L	membrane thickness in electrochemical permeation test	m
L	characteristic length of a capillary surface	m
m	flow rate in dropping mercury electrode	$kg\,s^{-1}$
M	atomic mass	$g\,mol^{-1}$
n	number of electrons transferred per molecule	dimensionless
nF	charge transferred per mole of species	$C\,mol^{-1}$
N	rotation rate	rpm
N	collection efficiency of a rotating ring disk electrode	dimensionless
N	total number of atoms in a particle	dimensionless
p_i	partial pressure of the ith components	atm
p_r	vapor pressure of a drop of radius r	atm
q_M	excess charge density on the metal surface	$C\,cm^{-2}$
q_S	excess charge density on the solution side of the interphase	$C\,cm^{-2}$
r	radius of the working electrode	m
r	the rate of change of the standard Gibbs free energy of adsorption with coverage	$J\,mol^{-1}$
r_{cyl}	radius of the RCE	m
R_F/R_{ct}	faradaic resistance/charge-transfer resistance	$\Omega\,m^2$
R_p	polarization resistance	$\Omega\,m^2$
R_S	uncompensated solution resistance	$\Omega\,m^2$
Re	Reynolds number	dimensionless

R_{film}	resistance of a nonconductive surface film	$\Omega\ m^2$
t	time	s
t	drop time in polarography	s
$t_{lag\_63\%}$	time during charging transient when the current density reaches 63% of the final steady-state current density	s
t_p	pulse duration in normal-pulse polarography	s
t^*	dimensionless pulse time in electroplating	dimensionless
T	thermodynamic (absolute) temperature	K
U	bond energy	$J\ mol^{-1}$
U_{cyl}	linear surface (peripheral) velocity of RCE	$cm\ s^{-1}$
U_{hyd}	hydration energy of ions in solution	eV
V	voltage drop	V
V_t	thermal voltage	V
v	potential sweep rate	$V\ s^{-1}$
v	heterogeneous reaction rate	$mol\ m^{-2}\ s^{-1}$
v	characteristic velocity of rotating disk electrode	$m\ s^{-1}$
ν	stoichiometric coefficient	dimensionless
ν	number of occurrences of the rate-determining step in the electrode reaction	dimensionless
v_{eq}	reaction exchange rate at equilibrium	$mol\ m^{-2}\ s^{-1}$
Wa	Wagner number	dimensionless
w	mass of a corroded metal	g
x	distance	m
$X^{\ddagger}$	activated complex	
z	charge number, valence	dimensionless
$Z(\omega)$	impedance	Ω
Z_W	Warburg impedance	Ω
α	transfer coefficient	dimensionless
α_a, α_c	anodic/cathodic transfer coefficient	dimensionless
β	symmetry factor	dimensionless
β_a, β_c	anodic/cathodic symmetry factor	dimensionless
γ	surface energy	$J\ m^{-2}$
γ_i	the activity coefficient of substance i	dimensionless
$\gamma_\pm$	mean activity coefficient	dimensionless
Γ	surface excess	$mol\ m^{-2}$
Γ	surface concentration	$mol\ cm^{-2}$
Γ_{max}	maximum surface concentration	$mol\ cm^{-2}$
δ	thickness of the Nernst diffusion layer	m
${}^{\alpha}\Delta^{\beta}\phi$	the difference between the inner potentials between two different phases, α and β	V
∇	del, the gradient operator	dimensionless

ε	absolute permittivity (dielectric constant) of the medium	$F\ m^{-1}$
ε_r	relative permittivity of the medium	dimensionless
ζ	zeta potential	V
η	overpotential	V
η	dynamic (absolute) viscosity	$kg\ m^{-1}\ s^{-1}$
η_a	anodic overpotential	V
η_c	cathodic overpotential	V
η_{ac}	activation overpotential	V
η_{conc}	concentration overpotential	V
η_{jR_s}	resistance (Ohmic) overpotential	V
θ	opening angle of a rotating cone electrode	degree (°)
θ	partial/fractional surface coverage	dimensionless
κ	specific conductivity of the solution	$S\ cm^{-1}$
κ	reciprocal Debye length	m^{-1}
λ	solvent reorganization energy	$J\ mol^{-1}$
λ_c	capillary length	m
μ_i	chemical potential of substance i	$J\ mol^{-1}$
μ^0	standard chemical potential	$J\ mol^{-1}$
$\overline{\mu}$	electrochemical potential	$J\ mol^{-1}$
Π	two-dimensional surface pressure	$J\ m^{-2}$
ρ	density	$kg\ m^{-3}$
ρ	reaction order	dimensionless
ρ	specific resistivity of the solution	Ω cm
σ	stored charge density	$C\ m^{-2}$
σ	surface tension	$N\ m^{-1}$
τ_c	time constant for the parallel combination of a capacitor and a resistor	s
τ_d	characteristic time constant for the diffusion process	s
τ_{pp}	total pulse time	s
υ	kinematic viscosity	$m^2\ s^{-1}$
ϕ	angle between the metal and the insulator	rad
ϕ^α	the inner potential of phase α	V
$\phi^\ddagger$	potential of the activated complex	V
ϕ_x	the potential at a distance x from the surface of the metal	V
ϕ_M	the electrode potential	V
ϕ_S	the potential in the bulk of the solution	V
$\Delta\phi$	the potential difference across the interface	V
$\Delta\phi_{rev}$	the value of $\Delta\phi$ at the reversible potential	V
χ	dimensionless rate constant	dimensionless
ω	angular velocity	$rad\ s^{-1}$

Abbreviations

AC	alternating current
ACD	anomalous codeposition
AES	Auger electron spectroscopy
AFC	alkaline fuel cell
AFM	atomic force microscope
AO	atomic oxygen
BET	Brunauer–Emmett–Teller
BEV	battery-electric vehicle
CE	counter electrode
CI	corrosion intensity
CMOS	complementary metal–oxide–semiconductor
CNT	carbon nanotube
CP	cathodic protection
CPE	constant phase element
CPR	corrosion penetration rate
CR	corrosion rate
CV	cyclic voltammetry
CVD	chemical vapor deposition
DC	direct current
DME	dropping mercury electrode
DMFC	direct methanol fuel cell
DPN	dip-pen nanolithography
DPP	differential-pulse polarography
EC	electrochemical capacitor
EcP	electrochemical printing
EDL	electrical double layer
EDLC	electrochemical double-layer capacitor
E-DPN	electrochemical dip-pen nanolithography
EDS	energy dispersive spectroscopy
EFAB	electrochemical fabrication
EIC	environmentally induced cracking
EIS	electrochemical impedance spectroscopy
EMF	electromotive force
EN	electrochemical noise
EPD	electrophoretic deposition
EQCM	electrochemical quartz crystal microbalance
ESC	environmental stress cracking
EV	electric vehicle
EW	equivalent weight
FC	fuel cell
FE	faradaic efficiency
FIB	focused ion beam
FPN	fountain pen nanofabrication
FRA	frequency response analyzer
FTIR	Fourier-transform infrared

HDME	hanging dropping mercury electrode
HE	hydrogen embrittlement
HER	hydrogen evolution reaction
HEV	hybrid electric vehicle
ICCP	impressed-current cathodic protection
ICE	internal-combustion engine
IHP	inner Helmholtz plane
LECD	localized electrochemical deposition
LIB	lithium-ion battery
LPR	linear polarization resistance
LSV	linear sweep voltammetry
MCED	meniscus-confined electrodeposition
MIC	microbiological corrosion
NHE	normal hydrogen electrode
NP	nanoparticle
NPP	normal-pulse polarography
OCP	open-circuit potential
OER	oxygen evolution reaction
OHP	outer Helmholtz plane
OPD	overpotential deposition
ORR	oxygen reduction reaction
PAFC	phosphoric acid fuel cell
PEM	polymer electrolyte membrane
PHEV	plug-in hybrid electric vehicle
QCM	quartz crystal microbalance
RConeE	rotating cone electrode
RCylE	rotating cylinder electrode
RDE	rotating disk electrode
RDS	rate-determining step
RE	reference electrode
RHE	reversible hydrogen electrode
RRDE	rotating ring disk electrode
SCC	stress corrosion cracking
SCE	saturated calomel electrode
SDME	static dropping mercury electrode
SECM	scanning electrochemical microscopy
SEI	solid/electrolyte interphase
SEM	scanning electron microscope
SGC	Stern–Geary coefficient
SHE	standard hydrogen electrode
SIMS	secondary ion mass spectrometry
SPE	solid polymer electrolyte
STM	scanning tunneling microscopy
SWP	square-wave polarography
UME	ultramicroelectrode
UPD	underpotential deposition
VCI	volatile corrosion inhibitor

WE	working electrode
XPS	X-ray photoelectron spectroscopy
XRD	X-ray diffraction
ZRA	zero-resistance ammeter

Useful Units and Conversions

Å	Angstrom	$1\ Å = 10^{-10}\ m = 10^{-8}\ cm = 10^{-4}\ \mu m = 10^{-1}\ nm$
Ah	Ampere-hour	$1\ Ah = 3600\ C$
C	Coulomb	$1\ C = 1\ A\ s$
cal	calorie	$1\ cal = 4.1868\ J$
dm	decimeter	$10\ dm = 1\ m$
Eq	equivalent	1 Eq will neutralize 1 mol of H^+ or $(OH)^-$ ions
erg	erg	$1\ erg = 10^{-7}\ J$
eV	electron volt	$1\ eV = 1.60218 \times 10^{-19}\ J$
F	Farad	$F \equiv C\ V^{-1} = A^2\ s^4\ kg^{-1}\ m^{-2}$
Hz	Hertz	$1\ Hz \equiv s^{-1}$
J	Joule	$J \equiv N\ m = kg\ m^2\ s^{-2}$
kWh	kilowatt hour	$1\ kW\ h = 3600\ kJ$
L	liter	$1\ L = 1 \times 10^3\ cm^3 = 1\ dm^3$
M	molal	$mol\ kg^{-1}$
mil	mil	$1\ mil = 0.001\ in.$
mpy	mils per year	$1\ mpy = 25.4\ \mu m\ y^{-1}$
M	molar	$mol\ L^{-1}$
N	Newton	$N \equiv kg\ m\ s^{-2}$
rad	radian	$1\ rad = 360/2\pi = 57.2958°$
S	Siemens	$1\ S \equiv 1\ A\ V^{-1} = kg^{-1}\ m^{-2}\ s^3\ A^2$
V	Volt	$V \equiv W\ A^{-1} = J\ C^{-1} = kg\ m^2\ A^{-1}\ s^{-3}$
W	Watt	$W \equiv J\ s^{-1} = kg\ m^2\ s^{-3}\ A^{-1}$
°F	degrees Fahrenheit	$°C = 5/9 \cdot (°F - 32)$
°K	degrees Kelvin	$°K = °C + 273.15$
Ω	ohm	$1\ \Omega \equiv 1\ V\ A^{-1} = kg\ m^2\ s^{-3}\ A^{-2}$

Physical Constants[1]

c	speed of light in vacuum	$2.99792 \times 10^8\ m\ s^{-1}$
e	elementary charge	$1.60218 \times 10^{-19}\ C$
F	Faraday constant	$9.64853 \times 10^4\ C\ mol^{-1} = 23\,060\ cal\ mol^{-1}\ V^{-1}$
g	standard acceleration of gravity	$9.80665\ m\ s^{-2}$

1 Taken from the CODATA Internationally Recommended 2014 Values of the Fundamental Physical Constants (http://physics.nist.gov/cuu/Constants/, last accessed July 12th, 2016).

h	Planck constant	6.62607×10^{-34} J s
k	Boltzmann constant	1.38065×10^{-23} J K^{-1}
K_w	equilibrium constant of water	1.008×10^{-14} mol^2 L^{-2} at 25 °C
N_A	Avogadro's number	6.02214×10^{23} mol^{-1}
R	molar gas constant	8.31447 J mol^{-1} K^{-1}
v_{tr}	transverse velocity of sound	3.34×10^4 m s^{-1} in AT-quartz
ε_0	permittivity of free space (electric constant)	8.85419×10^{-12} F m^{-1}
μ_q	shear modulus of quartz	2.947×10^{11} g cm^{-1} s^{-2}
ρ_q	density of quartz	2.648 g cm^{-3}

Potentials of Reference Electrodes in Aqueous Solutions at 25 °C

Common name	Electrode	V versus SHE	Notes
Mercury/mercurous sulfate (MMS)	Hg/Hg_2SO_4/0.5 M K_2SO_4	+0.680	Useful for avoiding chloride contamination of the test solution
Mercury/mercurous sulfate electrode (MSE)	Hg/Hg_2SO_4/ saturated K_2SO_4	+0.640	Useful for avoiding chloride contamination of the test solution
Calomel	Hg/Hg_2Cl_2/0.1 M KCl	+0.336	Better temperature stability than SCE
Copper/copper sulfate electrode (CSE)	Cu/saturated $CuSO_4$	+0.316	Very robust, commonly used for cathodic protection
Normal calomel electrode (NCE)	Hg/Hg_2Cl_2/1 M KCl	+0.280	Better temperature stability than SCE
Saturated calomel electrode (SCE)	Hg/Hg_2Cl_2/ saturated KCl	+0.241	The most common electrode in laboratory. Use of mercury introduces safety hazards. Potential decreases as the solubility of KCl increases at higher temperatures. Cannot be used above 50 °C
Saturated sodium calomel electrode (SSCE)	Hg/Hg_2Cl_2/ saturated NaCl	+0.236	
Saturated silver/silver chloride	Ag/AgCl/ saturated KCl	+0.197	Very easy to make, but light sensitive. Can be used up to 80–100 °C
Mercury/mercury oxide (MMO)	Hg/HgO/1 M NaOH	+0.140	Good for alkaline solutions
Standard hydrogen electrode (SHE)	H_2/H^+, $a_{H^+} = 1$ M, $f_{H_2(g)} = 1$ atm	0.000	Not to confuse with the normal hydrogen electrode (NHE) that implies $c_{H^+} = 1$ M, $E^0 \cong 0.000$ V

1

Introduction

1.1 General Considerations

1.1.1 The Transition from Electronic to Ionic Conduction

Physical electrochemistry is a branch of physical chemistry. If one were to describe the essence of electrode kinetics in one short phrase, it would be: "*the transition from electronic to ionic conduction*." Naturally, the transfer of charge is influenced by the potential. Conduction in the solution is ionic, whereas in the electrodes and the connecting wires it is electronic. The transition from one mode of conduction to the other requires charge transfer across the interface. The rate of this transfer is controlled by the catalytic properties of the surface and the adsorption on it, the concentration and the nature of the reacting species, and all other parameters that control the rate of heterogeneous chemical reactions. In addition, the potential plays an important role. This is not surprising, since charge transfer is involved, and this process can be accelerated or retarded by the magnitude and the sign of the potential applied across the interface.

Figure 1.1 shows a typical way that the current density would respond to the potential applied, for the case of a dilute (0.01 M) solution of HI in 1.0 M H_2SO_4. The reaction concerned is

$$2HI \rightarrow H_2 + I_2 \tag{1.1}$$

There is a lot of information in this simple equation, combined with Figure 1.1. The reaction does not occur spontaneously in the direction shown in Eq. (1.1). Electrical energy is needed to break up molecules of HI into H_2 and I_2. Electrical energy is being converted to chemical energy.

We note that Figure 1.1 has four regions. In the first region there is no current flowing, in other words, the resistance is infinite. It is important to understand that the current flowing here is *exactly zero*, not just very close to zero and too small to be detected experimentally. This follows directly from the first law of thermodynamics, the law of conservation of energy. The minimal electrical energy required to break up the molecules of HI in the particular reaction in Eq. (1.1) is 0.59 V. Above this threshold, the current rises exponentially with potential, along line a. The rate of the reaction is controlled by the rate of charge transfer j_{ct}. Eventually, the mass-transport limitation takes over, as shown by the horizontal part of line b shown in Fig. 1.1. This is the mass-transport-limited

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

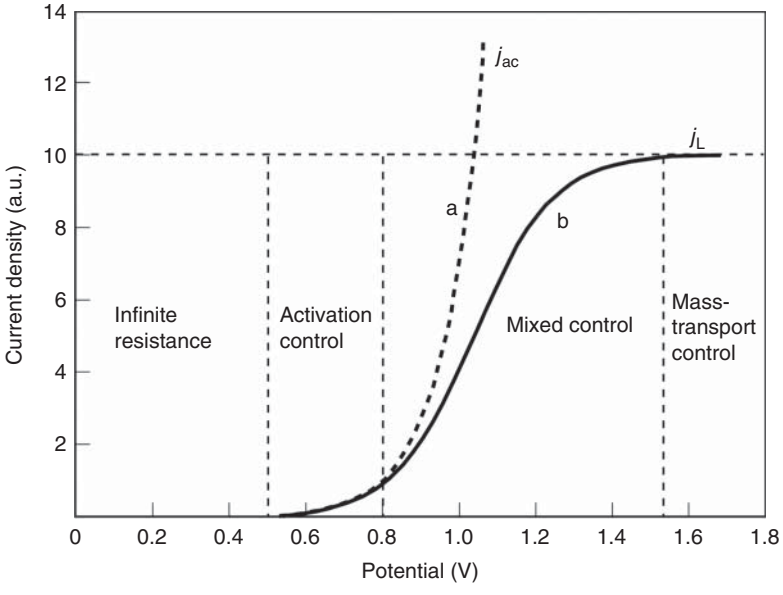


Figure 1.1 Schematic *j/E* plot for the electrolysis of a dilute (0.01 M) solution of HI in 1.0 M H_2SO_4, employing two Pt electrodes. The minimum potential for DC current flow is 0.59 V.

current density, j_L. In the third region, mixed control occurs; the current density is controlled both by activation and by mass transport.

1.1.2 The Resistance of the Interface can be Infinite

The potential is mostly measured with respect to a suitable reference electrode, and instead of the current, one refers to the current density *j*. Replacing the platinum electrodes with two copper electrodes and adding some $CuSO_4$ instead of HI changes the situation dramatically. Passing a current between the electrodes causes no net chemical change (copper is dissolved off the positive anode and is deposited on the negative cathode). In this case, current is observed as soon as a potential, small as it may be, is applied between the electrodes.

1.1.3 Mass-Transport Limitation

The rate of charge transfer can be greatly increased by increasing the potential, but charge can be transferred across the interface only over a very short distance (of the order of 1 nm). Another process is required to bring the reacting species close enough to the surface and to remove the species formed at the surface into the bulk of the solution. This process is called *mass transport*.

Mass transport and charge transfer are two consecutive processes. It is therefore always the slower of the two that determines the overall rate observed experimentally. When the potential applied is low, barely above its minimum value needed to pass a current, charge transfer is slow and one can ignore mass-transport limitation. The bottleneck is in transferring the charge across the

interface to the electroactive species, not in getting the species to the surface. At high potentials, charge transfer becomes the faster process and ceases to influence the overall rate. Increasing the potential further will increase the rate of charge transfer, but this will have no effect on the observed current density, which will be limited by mass transport. The result is a current density that is independent of potential, which is referred to as the *limiting current density*, j_L, as seen in Figure 1.1. For the observed current density j, one can write the simple equation

$$\frac{1}{j} = \frac{1}{j_{ac}} + \frac{1}{j_L} \tag{1.2}$$

Clearly, the smaller of the two currents is dominant. In a stirred solution, a steady state is reached when the concentration inside the diffusion layer varies linearly with distance. Under such conditions we can express the current density by

$$j = \frac{nFD(c_b - c_s)}{\delta} \tag{1.3}$$

where n is the number of electrons transferred, F is the Faraday constant ($96\,485\,\mathrm{C\,mol^{-1}}$), nF is the charge transferred per mole, D is the diffusion coefficient of the species ($\mathrm{cm^2\ s^{-1}}$), c_b is the concentration ($\mathrm{mol\,cm^{-3}}$) of the electroactive species in the bulk of the solution, c_s is its surface concentration ($\mathrm{mol\,cm^{-3}}$), and δ is the thickness of the Nernst diffusion layer (cm). Thus, the unit of current density is $\mathrm{A\,cm^{-2}}$. The current density reaches its mass-transport-limited value, j_L, when $c_s = 0$, thus

$$j_L = \frac{nFDc_b}{\delta} \tag{1.4}$$

The corresponding equation for the activation-controlled current density is

$$j_{ac} = nFkc_b \tag{1.5}$$

where the rate constant k is a function of the potential. From a comparison of the last two equations it is seen that the ratio D/δ in Eq. (1.4) has the same role as the rate constant k in Eq. (1.5), except that it is independent of potential. This ratio may be regarded as the specific rate of diffusion. From Eq. (1.4) it is also evident that the limiting current density increases as c_b is increased, the temperature is raised (thus, D increases), or the solution stirring is enhanced (thus, δ is decreased, as explained later).

Now, the essence of mass transport is the quantity δ. In certain favorable cases it has been calculated theoretically, in others it can only be determined experimentally. Sometimes, it is a function of time, while under different circumstances it is essentially constant during an experiment. Stirring the solution and transporting it toward, past, or through the electrode, all decrease the value of δ, hence, increase j_L. Moving the electrode (e.g. by rotation or vibration) has a similar effect. In quiescent solutions, δ increases linearly with $t^{1/2}$, hence j_L can be increased by taking measurements at short time.

In typical electrochemical measurements, the thickness of the Nernst diffusion layer attains values in the range of 10^{-3}–10^{-1} cm. Since in aqueous solutions at room temperature the diffusion coefficient of ions in aqueous solutions is on

the order of 10^{-5} cm^2 s^{-1}, this yields limiting current densities in the range of 0.01–1.0 mA cm^{-2} when $n = 1$ and the concentration of the electroactive species in solution is 10^{-6} mol cm^{-3}. The two most important things to notice in Eq. (1.4) are that: (i) the limiting current density is independent of the potential, and (ii) it depends linearly on the bulk concentration. A less obvious, but equally important, consequence of this equation is that j_L is independent of the kinetics of the reaction (i.e. of the nature of the surface and its catalytic activity). These characteristics make it an ideal tool for probing the concentration of species in solution. This is why most electroanalytical methods depend in one way or another on measurements of the mass-transport-limited current density.

1.1.4 The Capacitance at the Metal/Solution Interphase

When a metal is dipped in solution, a discontinuity is created. This affects both phases to some degree, so that their properties near the contact are somewhat different from their bulk properties. The exact position of the interface on the atomic scale is hard to define. "Where does the metal end?" we may ask. Is it the plane going through the centers of the outermost layer of atoms, is it one atomic radius farther out, or is it even farther out where the charge-density function of the free electrons in the metal has decayed to essentially zero? Fortunately, we do not need to know the position of this plane, for most purposes, when we discuss the properties of the interface.

One distinct property of the metal/solution interphase is a capacitance, called *the double-layer capacitance*, C_{dl}. It is a result of the charge separation between the two phases in contact. The double-layer capacitance observed depends on the structure of a very thin region near the interface, extending to about 1–10 nm, called *the double layer*. If the surface is rough, the double layer will follow its curvature down to atomic dimensions, and the capacitance measured under suitably chosen conditions is proportional to the *real* surface area of the electrode.

The double-layer capacitance is rather large, on the order of 10–30 μF cm^{-2}. This presents a serious limitation on our ability to study fast electrode reactions. Thus, a 10 μF capacitor coupled with a 10 Ω resistor yields a time constant $\tau_c = R \times C_{dl} = 0.1$ ms. It is possible to take measurements at shorter times by applying special techniques, but even so, the lower limit at present seems to be about 0.05 μs, seven orders of magnitude slower than that currently achievable in the gas phase.

The double-layer capacitance depends on the potential, the composition of the solution, the solvent, and the metal. It has been the subject of numerous investigations, some of which are discussed later (e.g. in Chapter 8).

1.2 Polarizable and Nonpolarizable Interfaces

1.2.1 Phenomenology

When a small current or potential is applied, the response is in many cases linear. The effective resistance may, however, vary over a wide range. When

this resistance is high, we have a polarizable interface, meaning that a small current generates a high potential across it (i.e. the interface is polarized to a large extent).

When the effective resistance is low, the interface is said to be nonpolarizable. In this case, a significant current can be passed with only minimal change of the potential across the interface. A nonpolarizable electrode is, in effect, a reversible electrode. The reversible potential is determined by the electrochemical reactions taking place and the composition of the solution, through the Nernst equation. For example, for a copper electrode in a solution containing $CuSO_4$ it is

$$E_{\text{rev}} = E^0 + \left(\frac{2.3RT}{nF}\right) \log\left(\frac{a_{Cu^{2+}}}{a_{Cu^0}}\right) \tag{1.6}$$

where $E^0 = +0.34$ V is the standard potential for the Cu^{2+}/Cu couple, on the scale of the Standard Hydrogen Electrode (SHE), and $a_{Cu^{2+}}$ is the activity of the cupric ions in solution. In aqueous solutions that are not very concentrated ($c_b \leq 1.0$ M), the error introduced when replacing the activity by the concentration is rather small and often considered negligible.

A good reference electrode, which exhibits a constant and stable potential, is always a reversible electrode. The inverse is not necessarily true. Not every reversible electrode is suitable as a reference electrode. For example, the correct thermodynamic reversible potential of a metal/metal-ion electrode may be hard to reproduce, because of impurities in the metal or complexing agents in the solution, even when the interface is highly nonpolarizable.

Polarizable interfaces behave differently. Their potential is not fixed by the solution composition, and it can be changed over a certain range depending on the metal and the composition of the solution in contact with it. For such a system, the potential may be viewed as an additional degree of freedom in the thermodynamic sense, as used in the Gibbs phase rule. To be sure, a so-called nonpolarizable interface can be polarized by passing a significant current through it. This, however, alters the concentration of both the reactant and the product *at the electrode surface* (without changing significantly their bulk concentrations). The potential developed across the interface will be in agreement with the Nernst equation as long as the concentrations used are the surface concentrations, which depend on the current passing across the interface.

We may summarize the difference between perfectly polarizable and perfectly nonpolarizable electrodes as follows. A perfectly polarizable electrode (e.g. Pt electrode) is an electrode in which no charge crosses the electrode/electrolyte interface when a current is applied. Such an electrode behaves like a capacitor. In contrast, a perfectly nonpolarizable electrode (e.g. Ag/AgCl/KCl reference electrode) is an electrode in which current passes freely across the interface, requiring no energy to make the transition.

1.2.2 The Equivalent Circuit Representation

We have already seen that the metal/solution interphase has some capacitance C_{dl} associated with it, as well as a (non-Ohmic) resistance R_F (also referred to as charge-transfer resistance, R_{ct}). Also, the solution has a finite resistance R_S that

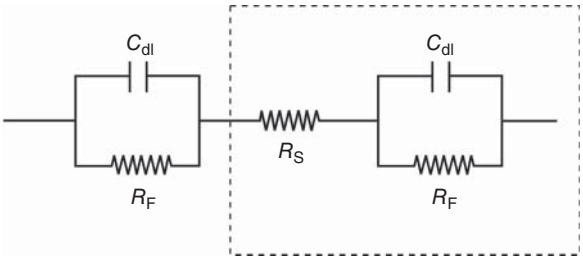


Figure 1.2 The equivalent circuit for a two-electrode cell. A single interface is usually represented by the elements inside the dashed rectangle (the Randles equivalent circuit). C_{dl}, R_F, and R_S represent the double-layer capacitance, the faradaic resistance, and the solution resistance, respectively.

must be taken into account. Thus, a cell with two electrodes can be represented by the equivalent circuit shown in Figure 1.2. Usually, one considers only the part of the circuit inside the dashed line, since the experiment is set up in such a way that only one of the electrodes is studied at a time. This part is known as the Randles equivalent circuit.

The equivalent circuit shown in Figure 1.2 represents a gross oversimplification, and interfaces rarely behave exactly like it. It does, nevertheless, help us gain some insight concerning the properties of the interface.

The combination of the double-layer capacitance and the faradaic resistance represents the interphase. How do we know that C_{dl} and R_F must be put in a parallel rather than in a series combination? Simply because we can observe a steady direct current flowing when the potential is high enough (above the minimum prescribed by thermodynamics (see Figure 1.1)). Also, when the resistance is effectively infinite under DC conditions, we can still have an AC signal going through (keep in mind that a capacitor does not transfer DC, only AC).

The equivalent circuit just described also makes it clear why conductivity measurements are routinely conducted by applying a small AC signal. If the appropriate frequency is chosen, the capacitive impedance associated with C_{dl} can be made negligible compared to the faradaic resistance, R_F, which is thus effectively shorted, leaving the solution resistance R_S as the only measured quantity.

The equivalent circuit shown in Figure 1.2 is meant to represent only the simplest situation. It does not take into account factors such as mass transport, heterogeneity of the surface, and the occurrence of reaction intermediates absorbed on it. Some of these factors are discussed later. Even in the simplest cases, in which this circuit does represent the response of the interphase to an electrical perturbation reasonably well, one should bear in mind that both C_{dl} and R_F depend on potential and, in fact, R_F depends on potential exponentially over a wide range, as will be discussed later.

The difference between polarizable and nonpolarizable interfaces can be easily understood in terms of this equivalent circuit. A high value of R_F is associated with a polarizable interface, whereas a low value of R_F represents a nonpolarizable interface.

Further Reading

1 Bard, A.J., Inzelt, G., and Scholz, F. (ed.) (2012). *Electrochemical Dictionary*. Berlin, Heidelberg: Springer-Verlag.
2 Zoski, C.G. (ed.) (2007). *Handbook of Electrochemistry*. Amsterdam, The Netherlands: Elsevier.
3 Calvo, E.J. (ed.) (2007). *Encyclopedia of Electrochemistry*, Interfacial Kinetics and Mass Transport, vol. 2. Weinheim: Wiley-VCH.
4 Bagotsky, V.S. (2006). *Fundamentals of Electrochemistry*, 2nd edn. New Jersey: Wiley.
5 Newman, J. and Thomas-Alyea, K.E. (2004). *Electrochemical Systems*, 3rd edn. New Jersey: Wiley.
6 Bockris, J.O'M., Reddy, A.K.N., and Gamboa-Aldeco, M. (2000). *Modern Electrochemistry*, Fundamental of Electrodics, 2nd edn, vol. 2A. New York: Kluwer Academic Publishers.
7 Randles, J.E.B. (1947). Kinetics of rapid electrode reactions. *Discuss. Faraday Soc.* 1: 11–19.

2

The Potentials of Phases

2.1 The Driving Force

2.1.1 Definition of the Electrochemical Potential

Knowledge of the driving force is of utmost importance for the understanding of any system. The driving force determines the direction in which a chemical reaction can proceed spontaneously, as well as its position of equilibrium at which the driving force is zero along all coordinates. The driving force in chemistry is the gradient of the chemical potential, μ:

$$\text{Driving force} = -\nabla\mu \tag{2.1}$$

where ∇ denotes gradient.

Let us turn our attention now to processes involving charged species, in particular charge-transfer processes. We recall that the chemical potential relates to the activity of the species:

$$\mu_i = \mu_i^0 + RT \ \ln(a_i) \tag{2.2}$$

The activity is related to the concentration via the activity coefficient γ_i, which is itself a function of concentration:

$$a_i = \gamma_i c_i \tag{2.3}$$

It would seem that μ_i, as given in Eq. (2.2), does not account for the effect of the electrical field or its gradient, unless we include that implicitly in the activity coefficient.

It has been found to be more expedient to define a new thermodynamic function, *the electrochemical potential,* $\overline{\mu}$, which includes a specific term to account for the effect of potential on a charged species:

$$\overline{\mu}_i = \mu_i + z_i F\phi^{\alpha} \tag{2.4}$$

where ϕ^{α} is the *inner potential* of a phase α, and z_i is the charge number on the *i*th species. When charged species are involved, the driving force is the gradient of electrochemical potential along some coordinate:

$$\text{Driving force} = -\nabla\overline{\mu}_i \tag{2.5}$$

It follows that the chemical potential is a special case of the electrochemical potential, applicable for an uncharged species.

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

2.1.2 Separability of the Chemical and the Electrical Terms

It has been noted earlier that Eq. (2.4) is an attempt to separate chemical interactions, represented by μ_i, from electrical interactions, represented by the product of charge (per mole) z_iF and the potential in the phase ϕ^α at a point *x,y,z*. But, can such separation be made?

One recalls that the potential $\phi_{x,y,z}$ at some point (*x,y,z*) in space is defined as the energy required to bring a unit positive test charge from infinity to that point. This is fine as long as the charge is moved in free space or inside a homogeneous phase. But what happens when we try to determine the potential inside a phase with respect to a point at infinity in free space, or the difference in potential between points in two different phases? As the "test charge" crosses the boundary of a phase, it interacts with the molecules in that phase, and it is impossible to distinguish between the so-called "chemical" and "electrical" interactions in this region. We must conclude from the above considerations that, while the electrochemical potential $\overline{\mu}$ is a measurable quantity, its components μ and ϕ cannot be separately measured.

The potential ϕ^α in Eq. (2.4) is called the inner potential in a phase, identified by the superscript. For the same reason that it cannot be measured, one cannot measure the value of ${}^\alpha\Delta^\beta\phi$, the difference between the inner potentials in two different phases. The foregoing statement may seem odd, since we are accustomed to measuring potential differences, say, the potential difference (i.e. the voltage) between two terminals of a battery. However, to do this, we connect the two terminals of a suitable voltmeter with copper wires to the terminals of the battery. We are therefore measuring, in effect, the potential difference between two *identical phases*, which, as we shall show further, is possible.

Now, consider an attempt to measure the potential difference across the metal/solution interphase $\phi^M - \phi^S \equiv {}^M\Delta^S\phi$. Assume that the metal used is copper, connected with a copper wire to one of the terminals of a voltmeter. This terminal will then be at the potential ϕ^M. Now, to determine the potential of the solution phase ϕ^S, we would have to use a copper wire connected to the other terminal of the voltmeter, and dip it in the solution. This, however, creates a new metal/solution interphase, and the voltmeter would show the sum of two metal/solution potential differences. It is important to realize that this is not a technical limitation, which may be overcome as instrumentation is improved. In any "thought experiment," one's attempt to measure the value of $\Delta\phi$ at a single interphase necessarily creates at least one more interphase.

To end this section on a positive note, it should be pointed out that, while ${}^M\Delta^S\phi$ cannot be measured, changes in it, $\delta({}^M\Delta^S\phi)$, (caused, for example, by passing a current) can be readily determined. Indeed, it is this quantity that is measured when *j/E* plots are shown in electrochemistry.

If the two terms on the right-hand side of Eq. (2.4) cannot be separately measured, what is the point of using this equation? It turns out that in some special cases of great practical importance, this equation does lead to results that can be tested by experiment. The usefulness of the electrochemical potential, as defined by Eq. (2.4), is in distinguishing between "short-range" interactions, represented by μ_i, and "long-range" interactions, represented by $z_iF\phi$. The energy of

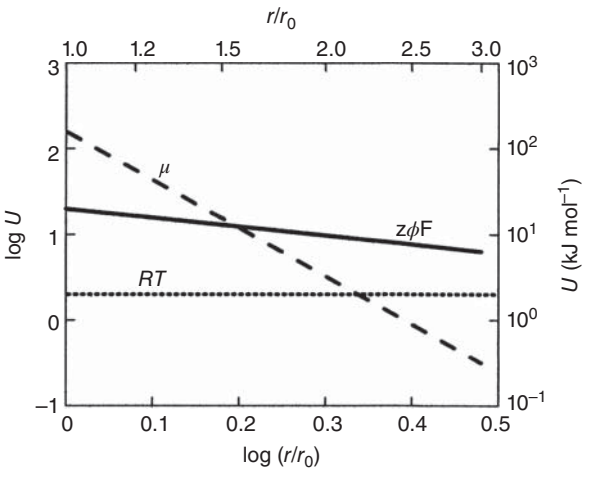


Figure 2.1 Variation of bonding energy with distance for short-range (covalent) interactions and long-range (electrostatic) interactions, marked by μ and $zF\phi$, respectively. The average thermal energy at 25 °C, RT, is given for comparison.

interaction for the former typically decays with r^{-6} while that for the latter decays with r^{-1}. This behavior is shown schematically in Figure 2.1.

In Figure 2.1, an initial value of 200 kJ mol^{-1} was chosen for the energy of the chemical bond, and the electrostatic energy was taken as 20 kJ mol^{-1} at the same distance. The lines for μ and $zF\phi$ intersect when the bond length is increased by 58%. The electrostatic energy becomes 10 times larger than the chemical energy when the bond length has been increased by a factor of 2.5. Thus, electrostatic interactions between charged species predominate everywhere, except very close to the boundary between two phases.

2.2 Two Cases of Special Interest

2.2.1 Equilibrium of a Species Between two Phases in Contact

Consider Eq. (2.4) for the case of a species at equilibrium in two different phases; for example, an electron in a copper wire and a nickel wire welded together. Since equilibrium is assumed, we can write

$$\overline{\mu}_{\mathrm{e}}^{\mathrm{Cu}} = \overline{\mu}_{\mathrm{e}}^{\mathrm{Ni}} \tag{2.6}$$

Combining with Eq. (2.4), we get

$$\mu_{\mathrm{e}}^{\mathrm{Cu}} - F\phi^{\mathrm{Cu}} = \mu_{\mathrm{e}}^{\mathrm{Ni}} - F\phi^{\mathrm{Ni}} \tag{2.7}$$

$$^{\mathrm{Cu}}\Delta^{\mathrm{Ni}}\mu_{\mathrm{e}}/F = {}^{\mathrm{Cu}}\Delta^{\mathrm{Ni}}\phi \tag{2.8}$$

In these and all following equations, the subscript represents the species and the superscript shows the phase in which it is.

The physical significance of Eq. (2.8) is that, at the contact between two dissimilar metals, a certain potential difference will develop, generated by the difference in the chemical potential of the electrons in the two metals. It might at first seem odd to have a potential drop inside a metal wire (even if it is made of two metals

welded together), unless a current is flowing. In the present case, however, one might consider ${}^{\mathrm{M}}\Delta^{\mathrm{S}}\phi$ as the potential difference needed to oppose the flow of electrons in the direction of decreasing chemical potential.

Consider a monovalent metal ion in solution, at equilibrium with the same ion in the crystal lattice. In this case

$$\mu^{\mathrm{M}}_{\mathrm{M}^+} + F\phi^{\mathrm{M}} = \mu^{\mathrm{S}}_{\mathrm{M}^+} + F\phi^{\mathrm{S}} \tag{2.9}$$

This leads to

$${}^{\mathrm{M}}\Delta^{\mathrm{S}}\mu_{\mathrm{M}^+}/F = -{}^{\mathrm{M}}\Delta^{\mathrm{S}}\phi \tag{2.10}$$

which is similar to Eq. (2.8). We recall that ${}^{\mathrm{M}}\Delta^{\mathrm{S}}\phi$ in these equations is not measurable, since any attempt to measure this quantity leads to the creation of at least one more interphase, with its own potential difference. Equation (2.10) does, however, lead to some very interesting conclusions, as we shall presently see. To do this, let us substitute the expression for chemical potential from Eq. (2.2) into Eq. (2.10):

$$\mu^{0,\mathrm{M}}_{\mathrm{M}^+} + RT \ \ln \ a^{\mathrm{M}}_{\mathrm{M}^+} - \mu^{0,\mathrm{S}}_{\mathrm{M}^+} - RT \ \ln \ a^{\mathrm{S}}_{\mathrm{M}^+} = -F({}^{\mathrm{M}}\Delta^{\mathrm{S}}\phi) \tag{2.11}$$

which can be written as

$${}^{\mathrm{M}}\Delta^{\mathrm{S}}\phi = ({}^{\mathrm{M}}\Delta^{\mathrm{S}}\phi)^0 - (RT/F)\ln\left(a^{\mathrm{M}}_{\mathrm{M}^+}/a^{\mathrm{S}}_{\mathrm{M}^+}\right) \tag{2.12}$$

where we have defined $({}^{\mathrm{M}}\Delta^{\mathrm{S}}\phi)^0$ as follows:

$$-({}^{\mathrm{M}}\Delta^{\mathrm{S}}\mu_{\mathrm{M}^+})^0/F = ({}^{\mathrm{M}}\Delta^{\mathrm{S}}\phi)^0 \tag{2.13}$$

Remembering that the activity of a species in a pure phase, $a^{\mathrm{M}}_{\mathrm{M}^+}$, is always defined as unity, we can rewrite Eq. (2.12) in a simplified form as follows:

$$\Delta\phi = \Delta\phi^0 + (RT/F)\ln(a^{\mathrm{S}}_{\mathrm{M}^+}) \tag{2.14}$$

which is a form of the Nernst equation, written for a single interphase. Combining two such equations, corresponding to two half-cells, leads to the Nernst equation in its usual form (i.e. in terms of the measurable cell potential).

2.2.2 Two Identical Phases not at Equilibrium

We turn our attention now to another special case of Eq. (2.4), in which the electrochemical potential of a species (an electron this time), $\overline{\mu}^{\mathrm{M}}_{\mathrm{e}}$, is considered in two identical phases that are not at equilibrium. In this case, we have

$$\overline{\mu}^{\mathrm{M}'}_{\mathrm{e}} = \mu^{\mathrm{M}'}_{\mathrm{e}} - F\phi^{\mathrm{M}'} \ \text{ and } \ \overline{\mu}^{\mathrm{M}''}_{\mathrm{e}} = \mu^{\mathrm{M}''}_{\mathrm{e}} - F\phi^{\mathrm{M}''} \tag{2.15}$$

and since

$$\mu^{\mathrm{M}'}_{\mathrm{e}} = \mu^{\mathrm{M}''}_{\mathrm{e}} \tag{2.16}$$

we obtain

$$\left({}^{\mathrm{M}'}\Delta^{\mathrm{M}''}\overline{\mu}_{\mathrm{e}}\right)/F = {}^{\mathrm{M}'}\Delta^{\mathrm{M}''}\phi \tag{2.17}$$

The quantity ${}^{\mathrm{M}'}\Delta^{\mathrm{M}''}\phi$ in Eq. (2.17) is the actual potential measured between two copper wires attached to the terminals of a battery. We note that potentials are

measured, as a rule, with a device that has a very high input resistance, which prevents the establishment of equilibrium between the electrons in its two terminals.

The important physical understanding that can be gained from Eq. (2.17) is that the potential difference measured is nothing but the difference in the electrochemical potentials of the electrons in the two terminals of the measuring instrument (divided by the Faraday constant, F, for consistency of units).

It is important to understand clearly the difference between the two special cases of Eq. (2.4). First, we discussed *equilibrium* between *dissimilar* phases. Then we discussed *non-equilibrium* between *identical* phases. The former led to

$$\left({}^{M'}\Delta^{M''}\overline{\mu}_e\right) = 0, \quad \text{hence,} \quad \left({}^{M'}\Delta^{M''}\mu_e\right)/F = \left({}^{M'}\Delta^{M''}\phi\right) \tag{2.18}$$

where $\left({}^{M'}\Delta^{M''}\phi\right)$ is not measurable. The latter led to

$$\left({}^{M'}\Delta^{M''}\mu_e\right) = 0, \quad \text{hence,} \quad \left({}^{M'}\Delta^{M''}\overline{\mu}_e\right)/F = \left({}^{M'}\Delta^{M''}\phi\right) \tag{2.19}$$

where $\left({}^{M'}\Delta^{M''}\phi\right)$ is the potential measured between the two (identical) terminals of a suitable voltmeter.

2.3 The Meaning of the Standard Hydrogen Electrode (SHE) Scale

Since it is impossible to determine experimentally the metal–solution potential difference at a single interphase, it is necessary to measure all such values against a commonly accepted reference. The reversible hydrogen electrode, operating under standard conditions ($T = 298.15$ K, pH = 0, $f_{H_2(g)} = 1$ atm) was chosen for this purpose, and its potential was assigned an arbitrary value of zero, $E^0 = 0.000$ V. The *standard hydrogen electrode* (SHE) is based on the reaction

$$2H^+(\text{aq, unit activity}) + 2e^-(\text{electrons in the metal}) \rightarrow H_2(\text{g, in solution, at unit activity}) \tag{2.20}$$

The standard chemical potentials of $H^+(aq)$[1] and $H_2(g)$ are both defined as zero in electrochemical equilibria calculations. One should note the difference between SHE and NHE (normal hydrogen electrode), the latter being rarely used nowadays. While the former implies that $a_{H^+} = 1$ M, $E^0 = 0.000$ V, the latter implies that $c_{H^+} = 1$ M, $E^0 \cong 0.000$ V. Note that the SHE is a theoretical electrode because one cannot have hydrogen ion with activity of 1 M in solution.

The SHE is called a *primary* reference electrode. The potential quoted for any redox couple on the SHE scale is then the potential measured in a cell made up of the desired redox couple in one half-cell (under the same standard conditions) and the SHE in the other half-cell. For example, taking the E^0 value of Cu/Cu^{2+} as +0.340 V versus SHE implies that this value will be measured versus a SHE.

1 The use of H^+ in all equations throughout this book is for simplicity. It should be recognized that a proton exists in a hydrated form such as H_3O^+ (hydronium ion), or even $H^+\cdot(H_2O)_n$, where $n > 1$.

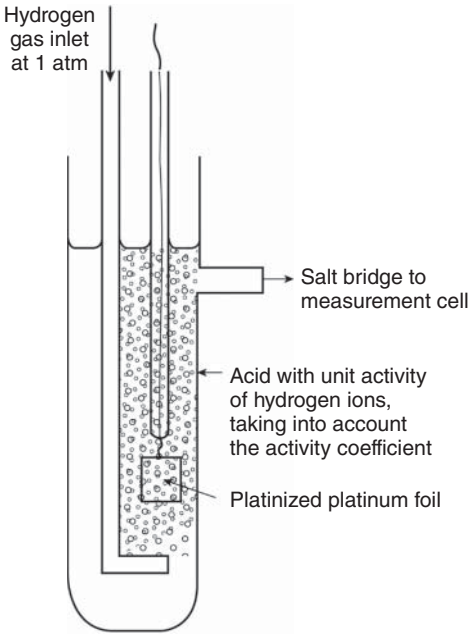


Figure 2.2 Schematics of the standard hydrogen electrode (SHE). Standard conditions: $T = 298.15$ K, pure hydrogen gas at $f = 1$ atm, pH = 0.

Figure 2.2 shows a schematic of the SHE electrode. The redox reaction occurs on a platinized platinum electrode, which is dipped in an acidic solution, and pure hydrogen gas is bubbled through it, keeping the solution saturated. Platinum serves as a catalyst for the hydrogen evolution reaction (HER); it does not dissolve in the solution, and therefore would not contaminate the solution. It is platinized to increase the total surface area, provide a surface material that adsorbs hydrogen well at its interphase, and improve the reaction kinetics.

The SHE has been very useful in defining the zero (reference) point in the *electromotive force (EMF) series*, relative to which the standard potentials of different half-cell reactions at standard conditions are arranged. However, it has rarely been used in the lab in real experiments because it is very difficult to assemble and also suffers from several drawbacks, as follows: (i) single ion activities cannot be determined, nor can they be adjusted independently; (ii) excess of H_2 bubbling out carries little HCl with it and, hence, the H^+ activity slightly changes; and (iii) if the solution contains any oxidizing agent, the H_2 electrode cannot be used.

What about sign convention? It has been internationally agreed that all potentials listed in the literature will refer to the reduction reaction. For the preceding two examples we can therefore write

$$Cu^{2+} + 2e_M^- \rightarrow Cu \quad E^0 = +0.340\ V_{SHE} \tag{2.21}$$

$$H_2 \rightarrow 2H^+ + 2e_M^- \quad E^0 = +0.000\ V_{SHE} \tag{2.22}$$

leading to

$$Cu^{2+} + H_2 \rightarrow Cu + 2H^+ \quad E^0 = +0.340\ V_{SHE} \tag{2.23}$$

and similarly

$$Fe^{2+} + H_2 \rightarrow Fe + 2H^+ \quad E^0 = -0.440\ V_{SHE} \tag{2.24}$$

A positive value of E^0 indicates that the reaction proceeds spontaneously in the direction shown, since for $E^0 > 0$ one gets $\Delta G^0 < 0$. We conclude from Eqs. (2.23) and (2.24) that copper ions in solution can be reduced by molecular hydrogen, while Fe^{2+} ions cannot. Looking at the same reactions proceeding in the opposite direction, we note that metallic copper cannot be dissolved in acid at pH = 0, while iron can.

If we were to combine copper with iron, we can combine Eq. (2.21) in its reduction form with the iron oxidation reaction

$$Fe \rightarrow Fe^{2+} + 2e^-_M \quad E^0 = +0.440\ V_{SHE} \tag{2.25}$$

to obtain the overall reaction

$$Fe + Cu^{2+} \rightarrow Fe^{2+} + Cu \quad E^0 = +0.780\ V_{SHE} \tag{2.26}$$

showing that this reaction proceeds spontaneously in the direction of dissolution of metallic iron and deposition of copper. This, in fact, is used as one of the industrial processes for copper recovery from ores. Note that Eq. (2.24) is written in the direction of reduction of Fe^{2+} ions, while Eq. (2.25) is written in the direction of oxidation of iron, hence the reversal of sign.

These observations are based on thermodynamic considerations alone. Thermodynamics can provide only the negative answers; it allows us to calculate and determine which reactions *will not* happen. We need kinetic information to determine what *will* happen at a rate that may be of practical interest, or at least at a rate that can be detected.

For the corrosion scientist it will be easy to remember that any metal for which E^0 is negative at a given pH is likely to corrode in acid, while those having a positive value of E^0 will not. This rule of thumb should not be taken as being exact, since in situations of practical interest the system is rarely, if ever, under standard conditions. Pipelines rarely carry 1.0 M acid, and metal structures are not, as a rule, in contact with a one-molar solution of their ions. For any specific system of known composition and pH, the reversible potential can readily be calculated from the Nernst equation, and the thermodynamic stability with respect to corrosion can be determined.

Further Reading

1 Eliaz, N. and Banks-Sills, L. (2008). Chemical potential, diffusion and stress – common confusions in nomenclature and units. *Corros. Rev.* 26: 87–103.

2 Newman, J. and Thomas-Alyea, K.E. (2004). *Electrochemical Systems*, 3rd edn. New Jersey, USA: Wiley.

3 Kaufman, M. (2002). *Principles of Thermodynamics*. New York, USA: Marcel Dekker.

4 Baierlein, R. (2001). The elusive chemical potential. *Am. J. Phys.* 69: 423–434.

5 Gavroglu, K. (1997). Philosophical issues in the history of chemistry. *Synthese* 111: 283–304.

6 Bratsch, S.G. (1989). Standard electrode potentials and temperature coefficients in water at 298.15 K. *J. Phys. Chem. Ref. Data* 18 (1): 1–21.

7 Ramette, R.W. (1987). Outmoded terminology: the normal hydrogen electrode. *J. Chem. Educ.* 64 (10): 885.

8 Bard, A.J., Parson, R., and Jordan, J. (1985). *Standard Potentials in Aqueous Solution*. New York, USA: Marcel Dekker.

9 Biegler, T. and Woods, R. (1973). The standard hydrogen electrode: a misrepresented concept. *J. Chem. Educ.* 50 (9): 604–605.

10 Hamer, W.J. (1968). Theoretical mean activity coefficients of strong electrolytes in aqueous solutions from 0 to 100°C. *National Standard Ref. Data Ser. (NSRDS) – NBS 24*. Washington, D.C.: National Bureau of Standards.

11 Garrels, R.M. and Christ, C.L. (1965). *Solutions, Minerals, and Equilibria*. New York, USA: Harper and Row.

12 Ives, D.J.G. and Janz, G.J. (ed.) (1961). *Reference Electrodes*. New York, USA: Academic.

13 Lewis, G.N., Randall, M., Pitzer, K.S., and Brewer, L. (1961). *Thermodynamics*, 2nd edn. New York, USA: McGraw Hill.

14 Gibbs, J.W. (1961). *The Scientific Papers of J. Willard Gibbs*, Thermodynamics, vol. I. New York, USA: Dover Publications Reprint.

15 Kielland, J. (1937). Individual activity coefficients of ions in aqueous solutions. *JACS* 59: 1675–1678.

16 Hückel, E. (1925). The theory of concentrated aqueous solutions of strong electrolytes. *Phys. Z.* 26: 93–147.

17 Debye, P. and Hückel, E. (1923). The theory of electrolytes. I. Lowering of freezing point and related phenomena. *Phys. Z.* 24: 185–206.

18 Lewis, G.N. (1907). Outline of a new system of thermodynamic chemistry. *Proc. Am. Acad. Arts Sci.* 43: 259–293.

19 Lewis, G.N. (1901). The law of physico-chemical change. *Proc. Am. Acad. Arts Sci.* 37: 49–69.

20 Gibbs, J.W. (1876/1878). On the equilibrium of heterogeneous substances. *Trans. Conn. Acad.* 3: 108–248, 343–524.

3

Fundamental Measurements in Electrochemistry

3.1 Measurement of Current and Potential

3.1.1 The Cell Voltage is the Sum of Several Potential Differences

The measured potential is the sum of several potential differences. When a current is made to flow through the cell, these potential differences are affected to different degrees, and the change in cell potential resulting from an applied current density, j, reflects the sum of all these changes

$$\delta E = E_{(j)} - E_{(j=0)} = \delta({}^{\mathrm{W}}\Delta^{\mathrm{S}}\phi) + jR_{\mathrm{S}} + \delta({}^{\mathrm{S}}\Delta^{\mathrm{C}}\phi) \tag{3.1}$$

where the superscripts W and C refer to the working and counter electrodes, respectively, and S stands for the solution phase. The voltage drop across the solution resistance, jR_{S}, can be calculated for certain simple and well-defined geometries. In most cases, however, it is measured and compensated for electronically. One is still left with δE representing the sum $\delta({}^{\mathrm{W}}\Delta^{\mathrm{S}}\phi) + \delta({}^{\mathrm{S}}\Delta^{\mathrm{C}}\phi)$, the changes of potential across the metal/solution interphase at the working and the counter electrodes, respectively. Several methods to overcome this problem and to relate δ E to $\delta({}^{W}\Delta^{S}\phi)$ have been devised and are discussed further.

3.1.2 Use of a Nonpolarizable Counter Electrode

If we combine the working electrode with a highly nonpolarizable counter electrode, the change of potential, $\delta({}^{\mathrm{S}}\Delta^{\mathrm{C}}\phi)$, at the counter electrode will be negligible compared to that at the working electrode, $\delta({}^{\mathrm{W}}\Delta^{\mathrm{S}}\phi)$, and practically all the change in potential observed will occur at the working electrode:

$$\delta E = \delta({}^{\mathrm{W}}\Delta^{\mathrm{S}}\phi) + \delta({}^{\mathrm{S}}\Delta^{\mathrm{C}}\phi) \cong \delta({}^{\mathrm{W}}\Delta^{\mathrm{S}}\phi) \tag{3.2}$$

This can be achieved either by using a highly reversible counter electrode or by making the counter electrode much larger than the working electrode. Since the same *total current* must flow through both electrodes, the *current density* at the counter electrode can be made much smaller than that of the working electrode. Now, the change of potential resulting from an applied current is determined by the current density, not the total current. Hence, $\delta({}^{\mathrm{S}}\Delta^{\mathrm{C}}\phi)$ can be made negligible compared to $\delta({}^{\mathrm{W}}\Delta^{\mathrm{S}}\phi)$, satisfying Eq. (3.2). This is so, even if the two electrodes are chemically identical and have the same inherent polarizability.

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

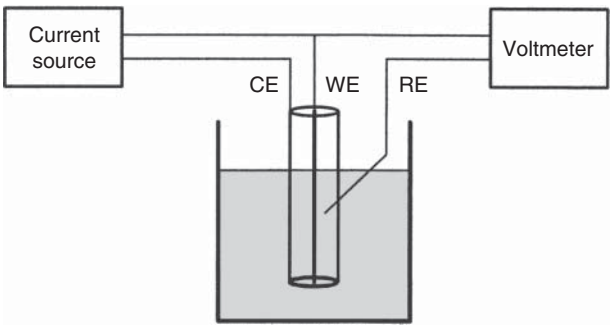


Figure 3.1 Schematic representation of a three-electrode setup, showing the working electrode (WE) connected to a reference electrode (RE) through a high-input-resistance voltmeter, and to the counter electrode (CE) through a low-input-resistance current source.

3.1.3 The Three-Electrode Setup

A better method of measuring changes in the metal-solution potential difference at the working electrode (which we shall refer to from now on as "changes in the potential of the working electrode") is to use a three-electrode system, shown schematically in Figure 3.1.

A variable current source is used to pass a current through the working and counter electrodes. Changes in the potential of the working electrode are measured versus a reference electrode, which carries practically no current. In this way, the polarizing current flows through one circuit (which includes the working and the counter electrodes), while the resulting change in potential is measured in a different circuit (consisting of the working and reference electrodes), through which the current is essentially zero. Since no current flows through the reference electrode, its potential can be considered to be constant, irrespective of the current passed through the working and counter electrodes. Thus, the measured change in potential (between working and reference electrodes) is truly equal to the change of potential of the working electrode. As stated in Chapter 2, although $\Delta\phi$ cannot be measured, its variation $\delta(\Delta\phi)$ can be readily determined.

It should be noted that during measurement in a three-electrode cell, the potential of the counter electrode might change substantially. This, however, does not in any way influence the measured potential of the working electrode with respect to the reference.

3.1.4 Residual jR_S Potential Drop in a Three-Electrode Cell

Regarded superficially, it might appear that making a current–potential measurement in a three-electrode cell eliminates the need to consider any correction for the jR_S potential drop in the solution, since there is practically no current flowing through the circuit used to determine the potential. Unfortunately, this is not quite true. The reference electrode (or the tip of the Luggin capillary leading to it) is situated somewhere between the working and counter electrodes. As a result, the potential it measures includes some part of the potential drop in the solution between these electrodes. This is called the *residual* jR_S potential drop. This is

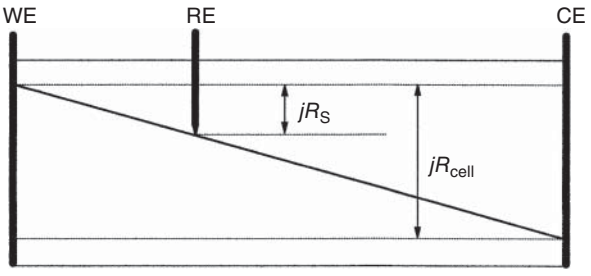


Figure 3.2 The residual jR_S potential drop between the working electrode and the tip of the Luggin capillary leading to the reference electrode. The total potential drop in the cell, between the working and the counter electrodes, is also shown.

shown schematically in Figure 3.2, where a cell having parallel-plate geometry is chosen for simplicity.

Placing the reference electrode near the working electrode can decrease, but not totally eliminate, the residual potential drop caused by residual solution resistance. The methods used to determine this potential drop are discussed later. It is important to understand that the jR_S potential drop *cannot* be determined by measuring the resistance between the terminals of the working and reference electrodes, since such a measurement includes resistive elements through which there is no flow of current during determination of the *j/E* relationship.

3.2 Cell Geometry and the Choice of the Reference Electrode

3.2.1 Types of Reference Electrodes

A type of reference electrodes commonly used consists of a metal (M), an insoluble salt of the same metal (MA), and a solution containing a soluble salt having the same anion (A), for example Ag | AgCl | KCl; H_2O. Here, the vertical lines represent separation between phases. This is a complete system, with the electrode and electrolyte enclosed in a small compartment and connected to the rest of the cell through a porous plug. The latter is designed to allow passage of ions, yet keep the flow of solution to a minimum. In this type of reference electrodes, the saturated solution of KCl is in contact with solid KCl. This ensures that its composition will be constant even if some of the water has evaporated over time, as long as the temperature is maintained constant, thus leading to a stable reference potential. Compatibility with the composition in the cell is, however, important. Thus, this reference electrode should not be used in a solution containing $HClO_4$, because $KClO_4$ could precipitate in the porous plug and isolate the reference electrode from the solution in the main compartment. Also, since chloride ions are strongly adsorbed on electrode surfaces and can hinder the formation of passive films in corrosion studies, a silver/silver chloride reference electrode should not be used, unless the test solution itself contains chloride ions, such as in sea water.

This electrode is an example of a so-called *secondary reference electrode*. A primary reference electrode comprises a metal at equilibrium with a solution containing its soluble salt, for example Cu | $CuSO_4$; H_2O. Secondary reference electrodes are usually preferred over primary reference electrodes because their potential is more stable and more reproducible.

Another class of reference electrodes, often called *indicator electrodes,* are electrodes in direct contact with the solution. The most common among these is the reversible standard hydrogen electrode, formed by bubbling hydrogen over a large-area-platinized Pt electrode in the test solution (see Figure 2.2). This electrode is reversible with respect to the hydronium ion, H_3O^+, serving, in effect, as a pH indicator electrode. Similarly, one could use a silver wire coated with AgCl in a chloride-containing test solution as a reversible $Ag\,|\,AgCl\,|\,Cl^-$ electrode, which responds to the concentration of Cl^- ions in solution, following the Nernst equation.

The advantage of indicator electrodes is that they always measure the reversible potential with respect to the ion being studied, regardless of its concentration in solution. Their disadvantage is that they must be prepared for each experiment and often end up being less stable and less reliable than commercial reference electrodes. Also, being in intimate contact with all ingredients in the test solution, they can be contaminated, either by impurities or by components of the test solution, such as additives for plating and corrosion inhibitors, which might lead to changes of the reference potential.

The Table in *Symbols and Abbreviations*, entitled "Potentials of Reference Electrodes in Aqueous Solutions at 25 °C," lists the most common reference electrodes, their potential versus the SHE, and some of their major characteristics and applications.

3.2.2 Use of an Auxiliary Reference Electrode for the Study of Fast Transients

One of the advantages of conducting measurements in a three-electrode configuration is that the resistance of the reference electrode should not affect the measured potential, since the current passing through it is extremely low. This allows us to use reference electrodes that are well separated from the main electrolyte compartment, thus minimizing the danger of mutual contamination. Typical values of the resistance of the reference electrode assembly may be about 10 kΩ. Combined with an input resistance of 10^{12} Ω of the voltmeter, the error in the measured potential is negligible. As is often the case, in practice the situation is more complicated, and it is advantageous to lower the resistance of the reference electrode. For one thing, the electrode and its connecting wires may act as antennas, picking up stray electromagnetic signals. The resulting noise in the measurement of potential will then increase with increasing resistance of the reference electrode. The problem is aggravated in the study of transients. The inherent reason for this is that there is a trade-off between response time and input impedance in all measuring instruments. While for steady-state measurements an input impedance of 10^{12} Ω is commonplace, fast oscilloscope and transient recorders may have an input impedance as low as 10^6 Ω.

Using an indicator-type reference electrode could alleviate this problem. When this is not possible (for example, because of chemical incompatibility), an "auxiliary reference electrode" can be used. This usually consists of a platinum wire placed near the working electrode. While the potential of such an electrode is not stable or well defined, it can be measured just before application of the transient,

and it can be safely assumed to be constant during the transient. The transient is then applied, with the platinum wire, which has a very small resistance, acting momentarily as the reference electrode. The potential of the Pt wire is measured before and after application of the transient, against a proper reference electrode.

3.2.3 Calculating the Uncompensated Solution Resistance for a few Simple Geometries

As a rule, the jR_S potential drop is measured, and a suitable correction is made, either directly during measurement or in the analysis of the data. When the geometry of the cell is simple, it is possible to calculate this quantity. Such calculations are important because they can yield clear criteria for the design of cells and for positioning the reference electrode with respect to the working electrode, as will be shown later.

3.2.3.1 Planar Configuration

The planar configuration is shown in Figure 3.2. The jR_S potential drop is given by

$$jR_S = j\frac{d}{\kappa} \tag{3.3}$$

where j is the current density ($A\,cm^{-2}$), R_S is the uncompensated solution resistance ($\Omega\,cm^2$), κ is the specific conductivity of the solution ($S\,cm^{-1}$), and d is the effective distance (cm) between the working electrode and the tip of the Luggin capillary connected to the reference electrode. The way to decrease jR_S is to increase the conductivity (e.g. by adding an inert supporting electrolyte), or by positioning the tip of the Luggin capillary close to the surface of the working electrode. However, the latter approach is not recommended, because it could result in local distortion of the current distribution, and the potential measured could deviate significantly from the metal–solution potential difference that exists over most of the electrode surface. We note that the jR_S is independent of the electrode area, because R_S is given in units of $\Omega\,cm^2$ and j has the units of $A\,cm^{-2}$.

3.2.3.2 Cylindrical Configuration

In a cylindrical configuration, the working electrode is positioned at the center of a cylindrical counter electrode. The equation relating jR_S to the distance d and the radius of the working electrode, r, is

$$jR_S = \frac{jr}{\kappa}\ln(1 + d/r) \tag{3.4}$$

It is interesting to consider two extreme cases of this equation. Close to the surface of the working electrode, where $d/r \ll 1$, the approximation

$$\ln(1 + d/r) \approx d/r \tag{3.5}$$

applies, yielding

$$jR_S = \frac{jd}{\kappa} \tag{3.6}$$

which brings us back to the equation for planar configuration. This should not be surprising, because any probe "looking" at a curved surface from a distance that is short compared to the radius of curvature responds to it as though it were flat. On the other hand, for large distances from the electrode surface, when $d/\mathrm{r} \gg 1$, one has

$$jR_S = \frac{jr}{\kappa} \ln \frac{d}{r} \tag{3.7}$$

There is little to be gained by decreasing the distance d (unless it can be reduced to well below the radius of the electrode), since jR_S changes logarithmically with the distance. On the other hand, we note that jR_S decreases in this case with decreasing radius of the working electrode. There is, therefore, a clear advantage in using a very fine wire in this type of measurement.

3.2.3.3 Spherical Symmetry

Next we consider the case of an electrode in the shape of a drop, located at the center of a spherical counter electrode. The potential drop across the solution resistance is expressed in this case by

$$jR_S = \left(\frac{jd}{\kappa}\right)\left(\frac{r}{r+d}\right) \tag{3.8}$$

For very short distances, $d/\mathrm{r} \ll 1$, Eq. (3.8) reverts to the equation for planar configuration, as for the cylindrical case. For large distances, however, the same equation yields

$$jR_S = \frac{jr}{\kappa} \tag{3.9}$$

What this equation shows is that for spherical symmetry, most of the jR_S potential drop occurs in the vicinity of the working electrode (within, say, 5 radii). Beyond that, it approaches a constant value, independent of distance. The variation of potential with distance is shown in Figure 3.3 for the three configurations just discussed.

It should be clear that the spherical configuration is the best in reducing the error that results from a residual jR_S potential drop, while the planar configuration is the worst. In spite of this, the cylindrical configuration is often used in research, because it is only a little worse than the spherical, but much better than the planar configuration, and is easier to implement experimentally than the spherical configuration.

The normalized resistance is shown for three radii of a spherical electrode in Figure 3.4. A similar behavior applies for the cylindrical configurations. In both cases, the error introduced by the potential drop associated with the uncompensated solution resistance can be reduced by reducing the radius of the electrode.

3.2.4 Positioning the Reference Electrode

In the design of an electrochemical cell, the position of the reference electrode has to fulfill two conflicting requirements. On the one hand, it should be far from the working electrode and well separated from the solution in the main cell

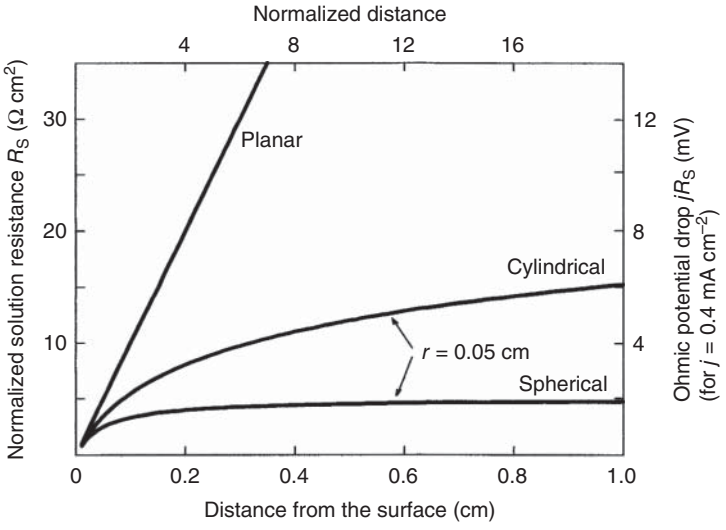


Figure 3.3 Uncompensated solution resistance and the corresponding potential drop, as a function of the distance from the electrode surface. Here, $\kappa = 0.01$ S cm^{-1}, $r = 0.05$ cm, the normalized distance is defined as d/r.

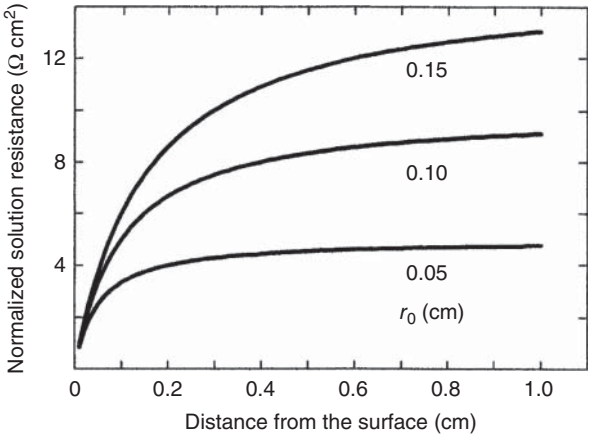


Figure 3.4 The normalized solution resistance at a spherical electrode, as a function of distance from the electrode surface, shown for different radii.

compartment, in order to reduce the possibility of mutual contamination. On the other hand, it should be as close as possible to the working electrode, in order to reduce the residual jR_S potential drop. This set of requirements is partially solved by the use of a Luggin capillary, shown in Figure 3.5.

The reference electrode compartment is separated from the rest of the solution by a porous plug and the tip of the capillary. Since the current to the reference electrode is essentially zero, the potential anywhere inside the capillary (and up to the reference electrode compartment) is the same as the potential at its outer rim, a distance d away from the working electrode.

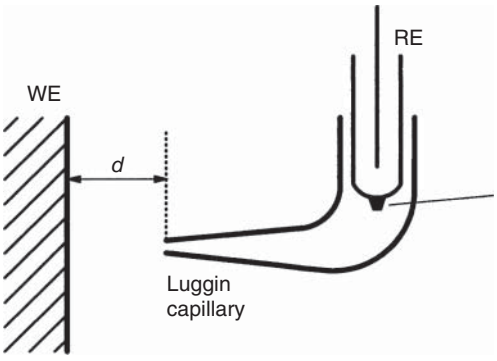


Figure 3.5 Schematic representation of a Luggin capillary. WE, working electrode; RE, reference electrode.

But how close should the tip of the capillary be to the surface of the electrode? Should it be in the middle of the electrode or near the edge? The former position measures a more typical value of the potential, not influenced by edge effects, but the body of the Luggin capillary may "cast a shadow" on the working electrode, disturbing the uniformity of the current density on its surface. The latter position is influenced by edge effects, but interferes less with the current flow. Many configurations have been suggested in the literature, and they all share the following drawbacks (albeit to varying degrees), which follow from the basic laws of electrostatics. Bringing the Luggin capillary close to the surface causes a nonuniformity of the current density in that area (usually, it is a decrease in local current density), caused by the existence of a nonconducting body, the glass capillary, in the path of the flow of current. If the Luggin capillary is small, this anomaly may not affect the total current to a significant extent. However, the potential is measured near the tip of the capillary, where the deviation of the current density from its average value is at maximum.

The configuration shown in Figure 3.5 was proposed by Luggin more than a century ago, when electronic instrumentation was in a very primitive stage, so that the Luggin capillary may have been the best way to reduce the residual solution resistance. Fortunately, modern instruments allow accurate measurement and dynamic compensation for the residual jR_S potential drop, obviating the need to use the Luggin capillary, except as a means for separating the solution in the reference electrode compartment from that in the working electrode compartment. If a Luggin capillary is used, it is therefore better to move it farther away from the working electrode (a distance of about five times the outer radius of the Luggin capillary is usually enough), in order to minimize the inhomogeneities of the current density distribution. In the case of cylindrical and spherical electrodes, this approach has the added advantage that the potential drop changes little with distance (at $d/r \geq 5$). Thus, the measured potential is not sensitive to the exact position of the Luggin capillary, and reproducibility is improved.

3.2.5 Edge Effects

Uniformity of the current density over the whole area of the electrode is important for the interpretation of current–potential data. We recall that the measured quantities are the total current and the potential at a certain point

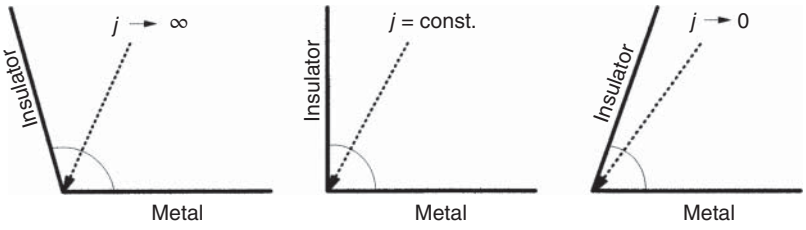


Figure 3.6 The angle ϕ between the metal and the insulator determines the current distribution near the edge of the electrode.

in solution, where the tip of the Luggin capillary is located. From this we can calculate the average current density, but not its local value. As for the potential, we have already noted that, unless the cell is properly designed, the potential measured might be grossly in error, if the reference electrode is located at a point where the local current density deviates significantly from its average value.

An interesting case to be discussed is the edge effect calculated for the point of contact between an electrode and the insulator in which it is cast. The equation describing the current density as a function of distance d from the point of contact is

$$j = Kd^{(\pi/2-\phi)} \tag{3.10}$$

in which, ϕ is the angle between the metal and the insulator. Three cases of interest are shown in Figure 3.6. It is easy to see that, according to Eq. (3.10), a uniform current density is expected only for a right angle ($\pi/2 = \phi$). For obtuse angles, the exponent in Eq. (3.10) is negative, hence, current density at the point of contact between the electrode and the insulator should approach infinity. For acute angles, the exponent is positive, and the current density approaches zero at the point of contact. This type of behavior is indeed observed experimentally and is well known in the plating industry.

A complete evaluation of the uniformity of current distribution should take into account the finite rate of the charge-transfer reaction that is taking place at the interface. When this is done, the current density becomes more uniform, and it neither declines to zero, nor does it increase to infinity at the edge, as implied above. It turns out that lower exchange current density for the metal deposition process and lower applied current density both enhance the uniformity of the current distribution.

Maintaining a uniform current distribution is of great importance in the electrochemical industry. In plating, it determines the uniformity of the thickness of the deposit; in electro-organic synthesis, it affects the uniformity of the products. A nonuniform current distribution can lead to the formation of undesired side products and to waste of energy in all areas of the electrochemical industry.

In studies of electrode kinetics it should be borne in mind that the fundamental equations used are derived with the tacit assumption of uniformity of current distribution. Hence, the analysis of the current–potential relationship is valid only when this assumption applies to a very good approximation.

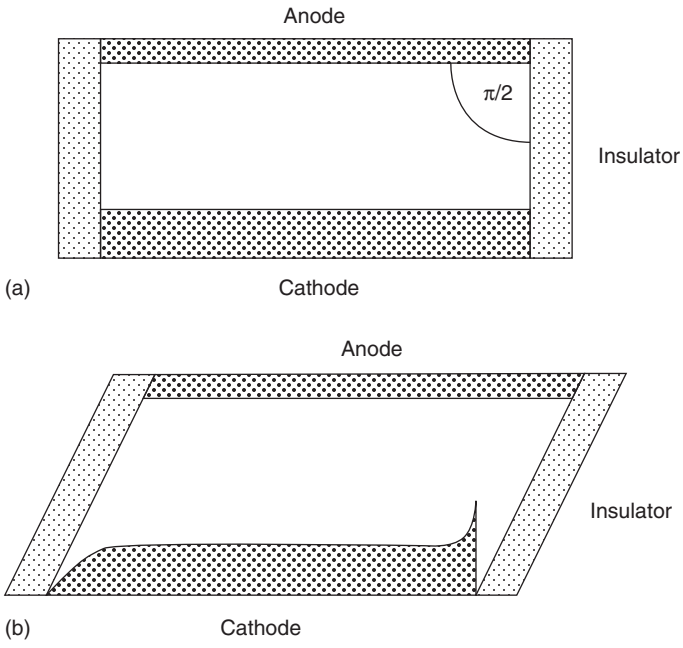


Figure 3.7 Coating thickness, indicative of current distribution, for different geometries. (a) Rectangular cell yielding uniform plating thickness. (b) Paralleloid-shaped geometry yielding near-zero thickness at the acute angle and high thickness at the obtuse angle.

The variation of the thickness of the electroplated layer is shown schematically in Figure 3.7. The deviations from uniformity at the edges for obtuse and acute angles are shown, in contrast to the uniformity expected when the contact between the insulator and the conductor is at right angles.

Further Reading

1 Bard, A.J. and Faulkner, L.R. (2001). *Electrochemical Methods: Fundamentals and Applications*, 2nd edn. New York, USA: Wiley.
2 Gileadi, E., Kirowa-Eisner, E., and Penciner, J. (1975). *Interfacial Electrochemistry – An Experimental Approach*. Addison-Wesley.
3 Newman, J.S. (1973). *Electrochemical Systems*. New Jersey, USA: Prentice-Hall.

4

Electrode Kinetics: Some Basic Concepts

4.1 Relating Electrode Kinetics to Chemical Kinetics

4.1.1 The Relation of Current Density to Reaction Rate

The current density is proportional to the rate of the heterogeneous reaction that takes place at the interface. The relationship

$$j = nFv \tag{4.1}$$

follows directly from dimensional analysis, since multiplying the units of the term nFv yields the unit of current density as shown in the following equation

$$(\mathrm{C\,mol^{-1}}) \times (\mathrm{mol\,cm^{-2}\,s^{-1}}) = (\mathrm{C\,cm^{-2}\,s^{-1}}) = \mathrm{A\,cm^{-2}} \tag{4.2}$$

Here, the symbols n, F, and v denote the number of electrons transferred per molecule, the Faraday constant, and the heterogeneous reaction rate, respectively. Substituting the appropriate numbers into Eq. (4.1) shows that the electrochemical reaction rate can be measured with very high sensitivity, without causing significant changes in the concentration of reactants or products in solution. This follows from the high sensitivity in measurement of current, hence also of charge. Thus, a current density of 1 $\mu\mathrm{A\,cm^{-2}}$, which can readily be measured accurately, corresponds to a very low reaction rate of about $10^{-11}\,\mathrm{mol\,cm^{-2}\,s^{-1}}$. Hence, one can ordinarily measure the rate of an electrode reaction for hundreds of seconds without causing a significant change in the concentrations of reactants or products in the bulk of the solution. As a result, most electrochemical reactions can be studied under *quasi-zero-order* conditions, since the change in concentration can be maintained to be negligible during measurement of the current.

Consider the following example, showing the high sensitivity that can be achieved by measurement of the current. The charge required to form a monolayer of hydrogen atoms adsorbed on platinum in the reaction

$$\mathrm{H_3O^+ + Pt + e_M^- \rightarrow Pt\text{-}H_{ads} + H_2O} \tag{4.3}$$

can be estimated as follows. The area taken up by a single platinum atom on the surface is on the order of $0.1\,\mathrm{nm^2}$. Hence, there are about 10^{15} platinum atoms per square centimeter of the metal surface. Considering that one electron is

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

discharged per platinum atom in Eq. (4.3), the total charge required to form a monolayer of adsorbed hydrogen atoms is

$$1.6 \times 10^{-19} \times 10^{15} = 1.6 \times 10^{-4} \ \text{C} \ \text{cm}^{-2} \tag{4.4}$$

This is an order-of-magnitude calculation. The correct number, obtained experimentally for hydrogen atoms adsorbed on platinum, is about 2.2×10^{-4} C cm^{-2}. Thus, it is necessary to pass a current of, say, 10 μA cm^{-2} for 22 seconds to form a monolayer. As far as the electrical measurement is concerned, one could easily measure a very small fraction of a monolayer, since 0.22×10^{-4} mC cm^{-2} equals about 2×10^{-9} mol cm^{-2} (for $n = 1$). In the case of hydrogen atoms, this amounts to 2 ng cm^{-2} of added weight. Thus, measurement of the current and time makes it possible to determine quantities of adsorbed hydrogen well below 1 ng cm^{-2}, which is better than the sensitivity achieved by the electrochemical quartz crystal microbalance (QCM), itself by far the most sensitive method of weighing small amounts of materials adsorbed from solution on a surface (see Chapter 16).

The same high sensitivity that makes measurements so convenient causes great difficulty in the electrolytic industry. It takes ~96,500 × n, equal to $C = 26.8 \times n$ A·h, to generate one mole of a product formed by electrolysis. A cylinder of compressed hydrogen commonly used in the laboratory contains about 0.5 kg of the gas. It would take a water electrolyzer running at a total current of 1 kA about 13 hours to produce this small amount of hydrogen! Incidentally, the same amount of electric charge would produce 35.5 times as much Cl_2 in the electrolysis of a solution of NaCl. The difference, of course, is the ratio of the atomic weight of Cl and H.

4.1.2 The Relation of Potential to Energy of Activation

In kinetics, the rate constant k can be written in the form

$$k = k_0 \exp\left(-\frac{\Delta G^{0\ddagger}}{RT}\right) \tag{4.5}$$

where $\Delta G^{0\ddagger}$ is the standard Gibbs free energy of activation and k_0 is a chemical rate constant (or pre-exponential factor). In Chapter 2 we have shown that for a charge-transfer process it is advantageous to separate the electrochemical potential into a chemical and an electrical term:

$$\overline{\mu}_i = \mu_i + z_i F \phi^{\alpha} \tag{2.4}$$

In the same way, we can express the change in the standard electrochemical Gibbs energy of a reaction as

$$\Delta \overline{G}^{0} = \Delta G^{0} \mp zF\Delta\phi \tag{4.6}$$

For the standard electrochemical Gibbs free energy of activation, we write

$$\Delta \overline{G}^{0\ddagger} = \Delta G^{0\ddagger} \pm \beta F \Delta\phi \tag{4.7}$$

Although Eqs. (4.6) and (4.7) look similar, the transition from one to the other is by no means trivial, and is the subject of detailed discussion. Here, we

shall limit ourselves to a brief discussion of two points. First, the dimensionless charge number (valence) of the substance, z, which appears in Eq. (4.6), has been dropped from Eq. (4.7), since it is tacitly assumed that electrode reactions occur by the transfer of one electron at a time.[1] Second, the parameter β, called the *symmetry factor*, has been introduced. By definition, β can assume values between zero and unity:

$$0 \leq \beta \leq 1 \tag{4.8}$$

and can be viewed as representing the fraction of the total change in the standard electrochemical Gibbs free energy of the reaction that is applied to its standard Gibbs free energy of activation:

$$\beta \equiv \frac{(\delta \Delta \overline{G}^{0\ddagger}/\delta \Delta \phi)}{(\delta \Delta \overline{G}^{0}/\delta \Delta \phi)} = \frac{\delta \Delta \overline{G}^{0\ddagger}}{\delta \Delta \overline{G}^{0}} \tag{4.9}$$

The Transition-State Theory, the symmetry factor, and the transfer coefficient are discussed in more detail in Chapters 5 and 6.

The important consequence of Eq. (4.7) for our present discussion is that the standard electrochemical Gibbs free energy of activation is a linear function of potential. The negative sign in Eq. (4.7) is applicable to anodic reactions, in which the rate is enhanced by increasing the potential in the positive direction, and the positive sign is applicable to cathodic reactions. In either case, the electrochemical Gibbs free energy of activation decreases when the potential is changed "in the right direction," namely positive for an anodic process and negative for a cathodic process. We can thus relate the anodic current density to the exponent of the potential:

$$j_a = nFv = nFk_0 c_b \exp\left(-\frac{\Delta \overline{G}^{0\ddagger}}{RT}\right) \exp\left(\frac{\beta \Delta \phi F}{RT}\right) \tag{4.10}$$

where v is the heterogeneous reaction rate (mol cm^{-2} s^{-1}) and k_0 is a chemical rate constant (cm s^{-1}).

Consider how the rate of a chemical or an electrochemical reaction can be changed. The best means of accelerating a chemical reaction is by increasing the temperature. This approach does have its limitations, however, because of possible decomposition of reactants or products, the range of stability of reaction vessels, the cost of the energy, and so on. Increasing the concentration of reactants is another means; employing a suitable catalyst (which decreases the value of $\Delta G^{0\ddagger}$) is a very effective method, most often used in industrial processes.

For an electrochemical reaction, all the above apply, but there is an additional, very powerful, means of controlling the rate of the reaction, namely by adjusting the potential. From Eq. (4.10) we can calculate the magnitude of this effect. To do this, we follow the common practice in electrochemistry and assume $\beta = 0.5$, although we shall see in Chapter 5 that this may not always be the case, and in any event, one should not take the symmetry factor to be $\beta = 0.500$ (i.e. *exactly* one half).

1 It has recently been shown that simultaneous transfer of two electrons is possible. Nevertheless, we consider here the transfer of one electron at a time, for simplicity.

For a change of potential of $\delta(\Delta\phi) = 1.0$ V at 25 °C, Eq. (4.10) yields a change of anodic current density by a factor of $\exp\left(\frac{0.5\times96{,}500}{8.314\times298}\right) = 3.5\times10^8$. The effect of temperature on the reaction rate depends on the enthalpy of activation. Taking a reasonable range of 40–80 kJ mol^{-1} for this quantity, we find that the rate of a reaction at 100 °C is larger than at 25 °C by a factor of 1.3×10^2–1.6×10^4. Thus, the effect of raising the temperature by 75 °C on the reaction rate is many orders of magnitude less than the effect of changing the potential by one volt.

4.1.3 Mass-Transport Limitation Versus Charge-Transfer Limitation

In Chapter 1 we gave a schematic presentation (see Figure 1.1) of the current–potential relationship in an electrolytic cell, from the region where no current density is flowing in spite of the applied potential, to the region where the current density rises exponentially with potential, following an equation such as Eq. (4.10), and through the limiting current density region, where it has a constant value that is determined only by the rate of mass transport to the electrode surface or away from it.

Mass-transport limitation is more often encountered in electrode kinetics than in any other field of chemical kinetics, because the activation-controlled charge-transfer rate can be accelerated (by applying a suitable potential) to the point where it is much higher than the rate of mass transport, and therefore no longer controls the observed current. From a research perspective, mass transport is an added complication to be either avoided or corrected quantitatively, in order to obtain the true kinetic parameters for the charge-transfer process. From an engineering perspective, mass transport is often the main factor determining the space-time yield of an electrochemical reactor, i.e. the amount of product that can be generated in a given reactor (or a given plant) per unit time and unit volume of the reactor. The task of the engineer in industry is then twofold: first to find better catalysts (or rather electro-catalysts), which accelerate the rate of charge transfer, and then to design cells with high rates of mass transport, to take full advantage of the improvement in catalytic activity.

Mass transport to the interphase can occur through three independent mechanisms: *migration, convection*, and *diffusion*. The driving force for migration is the electric field in solution. In setting up an experiment, one usually tries to eliminate this effect by adding a high concentration of supporting electrolyte (compared to the concentration of the electroactive species). The supporting electrolyte is chosen such that it will not take part in the charge-transfer process. It is added to reduce the electric field by increasing the conductivity in solution. In addition, most of the electricity is carried by the inert ions of the supporting electrolyte, so that the contribution of the electrical field to the rate of mass transport is reduced to a negligible level. In the following discussion, it is assumed that mass transport by migration has been essentially eliminated, unless we refer to it specifically. It should be borne in mind, however, that this is not always the case. Specifically, when films are formed on the surface of electrodes (such as anodic oxide films on aluminum and some other metals of the so-called solid/electrolyte interphase), mass transport through the film may depend exclusively on migration. Also, in many industrial processes, a high concentration of the electroactive

ions is used to allow high rates of production (e.g. Cl_2 from NaCl, Cu plated from acid bath containing H_2SO_4 and $CuSO_4$), and migration is viewed favorably, as an additional mode of mass transport.

Convective mass transport is caused by the movement of the solution as a whole. The driving force in this case is external energy, usually in the form of mechanical energy of stirring the solution, rotating the electrode, or pumping the liquid at, past, or through the electrode. Most efficient convective mass transport is achieved when the flow at the electrode surface is turbulent; yet, turbulent flow tends to be sensitive to minor changes in cell configuration, electrode area, the nature of the electrode surface and so on, often making it difficult to obtain reproducible results.

Mass transport by diffusion can be regarded as the last resort. When movement of the electroactive species is not promoted by the input of external energy, either electrical (migration) or mechanical (convection), diffusion takes over. The driving force in this case is the gradient in chemical potential caused by the gradient in concentration. It is a relatively slow process, with diffusion coefficients for small molecules in dilute aqueous solutions at room temperature being in the range of 2×10^{-6}–8×10^{-5} cm^2 s^{-1}.

4.1.4 The Thickness of the Nernst Diffusion Layer

The thickness of the Nernst diffusion layer, δ, was discussed in Chapter 1. This quantity is related to the mass-transport-limited current density through the equation

$$j_L = \frac{nFDc_b}{\delta} \tag{1.4}$$

For a diffusion-controlled process, δ is proportional to the square root of time:

$$\delta = (\pi Dt)^{1/2} \tag{4.11}$$

Hence, the limiting current density decreases gradually with time.

The concept of the diffusion-layer thickness can also be applied, however, when the main mode of mass transport is convection. In such cases, one assumes that there is a thin layer of liquid at the surface that is stationary, while the rest of the solution is stirred. Concentration profiles near the interphase, following the application of a potential step, are shown in Figure 4.1.

In Figure 4.1a, the concentration profile is presented in dimensionless form, as the dimensionless concentration, c/c_b, versus the dimensionless distance, $x/(4Dt)^{1/2}$. This is an elegant way of representing the results calculated from a function that depends on several parameters. Thus, this curve is independent of the initial concentration in solution, the diffusion coefficient of the reacting species, and time. A great deal of information is condensed into a single curve, which can then be used to calculate values of the concentration as a function of distance and time for any specified system.

But how far is one unit of dimensionless distance in real terms? That depends on the diffusion coefficient and on time. For $D = 1 \times 10^{-5}$ cm^2 s^{-1} it is 2.0 μm after 1 ms, and 63 μm after 1 s. Figure 4.1b shows concentration profiles calculated

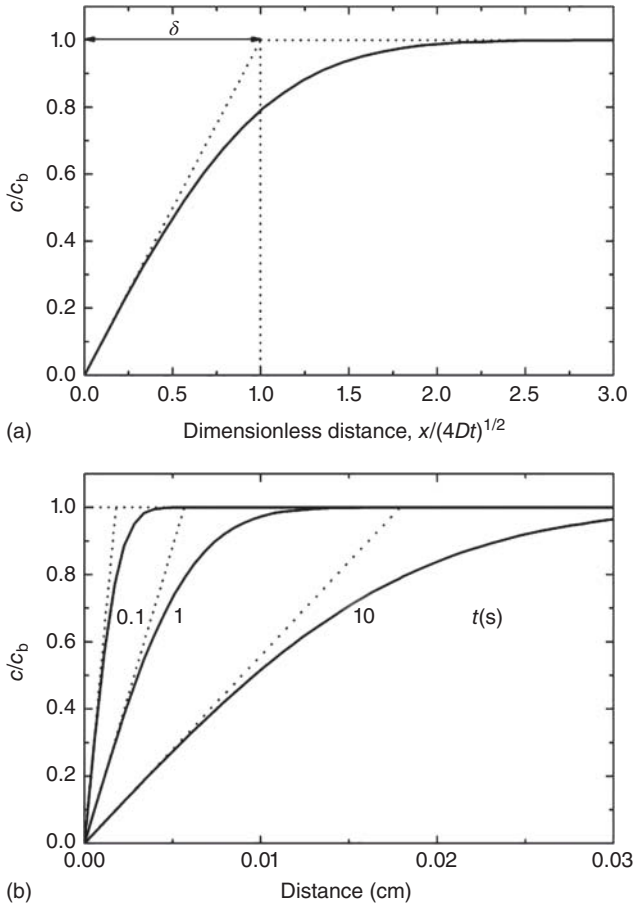


Figure 4.1 Calculated concentration profiles following a potential step. $D = 8 \times 10^{-6}$ cm^2 s^{-1}. (a) dimensionless coordinates, (b) distance in cm, with the time marked on each curve.

for the same system, but plotted versus the distance in centimeters. Each curve corresponds to a different time after application of the potential step, and the evolution of the concentration profile with time is presented. All the information contained in the curves in Figure 4.1b exists, of course, in the single curve shown in Figure 4.1a. This is both the advantage and the disadvantage of presenting the data in dimensionless form.

The current density flowing across the interphase is given by

$$j = -nFD(\mathrm{d}c/\mathrm{d}x)_{x=0} \tag{4.12}$$

It is determined by the gradient of concentration at the electrode surface, which decreases with time, as seen in Figure 4.1b.

In a stirred solution, a steady state is reached when the concentration at the metal/solution interphase is zero. Under such conditions we can express the current density by

$$j = nFD(c_\mathrm{b} - c_\mathrm{s})/\delta \tag{1.3}$$

It is evident that it can reach its mass transport limited value, j_L, when $c_s = 0$, namely

$$j_L = \frac{nFDc_b}{\delta} \tag{1.4}$$

Combining the last two equations yields

$$j/j_L = 1 - c_s/c_b \tag{4.13}$$

or

$$c_s/c_b = 1 - j/j_L \tag{4.14}$$

Since the reactant is consumed at the interphase, its surface concentration, c_s, is always smaller than the bulk concentration, c_b. The ratio between the two depends on the ratio j/j_L. For proper kinetic measurements, without the influence of mass transport, this ratio must be maintained below a chosen level, for example 0.01 or 0.05, depending on the accuracy desired. This can be done by taking measurements at very low current densities, by increasing j_L, or both. Alternatively, the concentration at the surface can be allowed to deviate significantly from its bulk value, as long as this deviation is taken into account quantitatively, by solving the appropriate equations for mass transport.

For a diffusion-controlled process, δ is a rough estimate of the distance over which molecules can diffuse in a given time. For $\delta = 50$ μm and $D = 8 \times 10^{-6}$ cm^2 s^{-1}, Eq. (4.11) yields a value of $t = 1$ second. During this time, some molecules will be found further away than 50 μm from the surface. Thus, δ should be considered as "the characteristic length" for diffusion. This kind of simple calculation allows us to estimate how long it would take, for example, for chloride ions from a Ag/AgCl/saturated KCl reference electrode to penetrate the test solution compartment through a side arm of given length, or how fast ions will diffuse into the pores of an anodized aluminum electrode, to be deposited there electrochemically.

4.2 Methods of Measurement

4.2.1 Potential Control Versus Current Control

An experiment in electrode kinetics usually consists of determining the current–potential relationship under a given set of fixed conditions. The measurement may then be repeated under a set of gradually changing conditions, to obtain the j/E plots as a function of temperature, concentration, etc.

In many cases, the experiment can be performed either by controlling the current externally and measuring the resulting changes in potential at the working electrode (i.e. between the working electrode and a suitable reference electrode), or by controlling the potential and measuring the resulting current. The former is referred to as a *galvanostatic*, and the latter is called a *potentiostatic* measurement. Admittedly, these commonly used terms are misnomers to some extent, since they include cases in which either E or j is changed in the course of

measurement, and the suffix "static" does not really apply. They serve, however, to identify the variable that is controlled externally and that which is measured. Even though galvanostatic and potentiostatic measurements may yield the same results in many cases, it is important to understand the inherent difference between them and to determine the advantages and disadvantages of each.

Controlling the potential in measurements of electrode kinetics is in many ways "the natural way" to conduct a measurement. It is similar to any other measurement in chemical kinetics in that all the experimental parameters (concentration, temperature, the nature of the catalyst, and in electrochemistry – also the potential) are controlled externally and the current density, which is proportional to the reaction rate, is measured (see Eq. (4.10)). Just as in chemical kinetics, we may repeat the experiment, changing the concentration of one of the components or the temperature. In electrochemistry we have the additional option of changing the potential.

The potential of an electrode with respect to a reference electrode represents its oxidizing or reducing power. In a mixture of, say, silver and copper ions (E^0 equals +0.799 V and +0.337 V versus SHE, respectively), one could set the potential at a value that ensures that only silver ions can be reduced, producing a pure deposit of this metal, even if we leave the system operating for a very long time. Similarly, it is possible, in principle at least, to determine the course of an electro-organic synthesis to produce one desired product or another, by prudent choice of the potential.

A galvanostatic measurement represents a different situation, unparalleled in chemical kinetics. Here, the rate of the reaction (i.e. the current density) is controlled externally, and the potential is allowed to assume a value appropriate to that rate.

The choice between galvanostatic and potentiostatic measurements depends on the circumstances. From the instrumentation point of view, galvanostats are much simpler than potentiostats. This is not only a matter of cost, but also a matter of performance. Thus, where it is desired to measure very low currents (e.g. on single microelectrodes), a battery with a variable resistor may be all that is needed to set up a low-noise galvanostat. At the other extreme, when large currents must be passed, for instance, in an industrial pilot plant for electrosynthesis, power supplies delivering controllable currents in the range of hundreds of amperes are readily available, whereas potentiostats of comparable output are either nonexistent or extremely expensive. In large-scale industrial processes, the current is usually the controlled parameter.

A potentiostat is inherently a more complex apparatus. It operates with a feedback loop, in which the potential between the *working* and the *reference* electrodes is compared to a value determined externally, the difference is amplified, and a current is passed between the *working* and *counter* electrodes, at such magnitude and sign as to reduce that difference of potential to a very small value, typically 1 mV or less.

Potentiostats did not become commercially available until the late 1950s. Earlier work was conducted either galvanostatically or potentiostatically, but with a two-electrode cell, in which one of the electrodes served as both the counter and the reference electrode. Because of their complexity, potentiostats

tend to have slower response times than galvanostats. It should be pointed out, though, that some of the limitations of potentiostats alluded to above are a matter of the past. With present-day electronic components (2018), it is possible to build home-made potentiostats, or to purchase commercial units, that make use of all the inherent advantages of potentiostatic measurements with few instrumental limitations.

4.2.2 The Need to Measure Fast Transients

So far, we have focused our attention on activation-controlled processes. Since the rate of such processes increases exponentially with potential, it is usually possible to drive them at a rate sufficiently high, so that mass transport becomes the limiting factor. For a measurement to be truly activation controlled, the current density must be small compared to the mass-transport-limited current density. The latter is given by

$$j_{\mathrm{L}} = \frac{nFDc_{\mathrm{b}}}{\delta} \tag{1.4}$$

where δ is the thickness of the Nernst diffusion layer. The activation-controlled current density, j_{ac}, is given by an equation of the form

$$j_{\mathrm{ac}} = nF\nu = nFk_0c_{\mathrm{b}} \exp\left(-\frac{\Delta\overline{G}^{0\ddagger}}{RT}\right) \exp\left(\frac{\beta\Delta\phi F}{RT}\right) \tag{4.10}$$

To maintain the ratio $j_{\mathrm{ac}}/j_{\mathrm{L}}$ small (say, <0.05) over a wide range of potentials, we would like to find ways of increasing j_{L} without changing j_{ac}. The best way to achieve this is to decrease the diffusion-layer thickness, δ. This can be done by improving the efficiency of stirring, or by taking measurements in unstirred solutions at short times. In quiescent solutions, diffusion is the only mode of mass transport, and the diffusion-layer thickness is given by

$$\delta = (\pi Dt)^{1/2} \tag{4.11}$$

Taking measurements at short times increases the value of j_{L}, allowing us to study j_{ac} over a wider range of potentials. Equation (4.11) is strictly applicable only in unstirred solutions, but as long as the value of δ calculated from it for a purely diffusion-controlled process is small compared to the thickness of the diffusion layer set up by stirring, the latter will have no effect on j_{L}.

Now that the importance of transient measurements in electrode kinetics has been discussed, it is of interest to consider the effect of the residual jR_{S} potential drop in solution on the analysis of such transients. Consider the equivalent circuit shown in Figure 4.2. It is important to understand that, although the potential that is applied in a potentiostatic measurement (or being measured in a galvanostatic measurement) is between points A and C, the potential driving the charge-transfer process is that between points B and C. The difference between these two potentials is given by the uncompensated potential drop

$$E_{\mathrm{AB}} = jR_{\mathrm{S}} \tag{4.15}$$

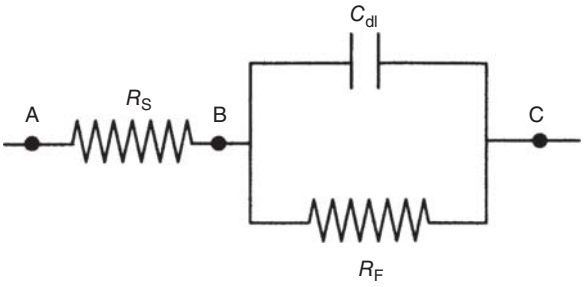


Figure 4.2 The equivalent circuit for an interphase, showing the difference between the measured potential E_{AC} and the potential actually applied across the interphase, E_{BC}, which drives the charge-transfer process.

Consider the effect of the uncompensated solution resistance on a cyclic voltammetry experiment, in which a cyclic linear sweep is applied between the working and the reference electrodes. A typical current/time response is shown in Figure 4.3. The corresponding variations of the potential E_{AC} applied externally and the potential E_{BC} actually imposed on the interphase are shown schematically in Figure 4.4.

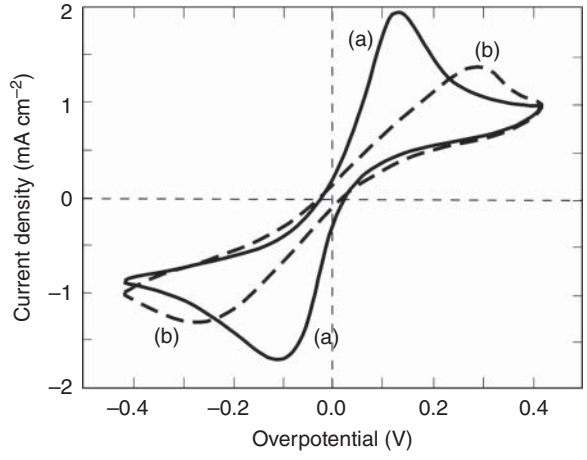


Figure 4.3 Variation of current density with time during linear potential sweep measurement in a poorly conducting electrolyte. (a) with dynamic $E_{AB} = jR_S$ compensation, and (b) without jR_S compensation. Conducted in a solution of 5 mM quinhydrone and 1 mM H_2SO_4, $v = 75$ mV s^{-1}.

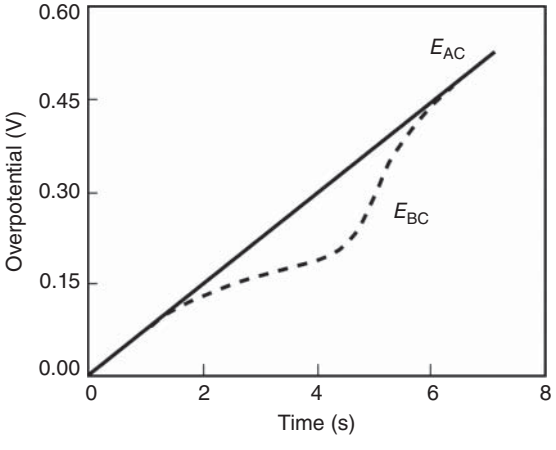


Figure 4.4 Schematic representation of the variation of potential with time during linear potential sweep. Solid line: the applied potential, E_{AC}. Dashed line: the potential actually imposed across the interphase, E_{BC}.

While the potentiostat delivers a constant rate of change of potential with time,

$$v = \mathrm{d}E/\mathrm{d}t = \text{constant} \tag{4.16}$$

the actual rate of change of potential imposed on the interphase (dashed line in Figure 4.4) varies during the transient. This can cause a major shift in the position of the peak current density along the potential axis and modify the numerical value of the peak current, j_p, as seen in Figure 4.3. Analysis of the *j/E* plots is based on the assumption that the potential applied across the interphase varies linearly with time. If the potential drop across the uncompensated solution resistance is significant, it could lead to gross errors in the interpretation of the results. It may thus be concluded that potentiostatic transients can be analyzed correctly only if the value of jR_S is negligible throughout the transient, or if it has been compensated for in real time.

4.2.3 Polarography and the Dropping Mercury Electrode (DME)

Polarography is one of the earliest electrochemical techniques, which had a major impact on the development of electroanalytical chemistry. Developed by Heyrovsky in the 1920s (and recognized by the Nobel Prize he received in 1959), it was implemented by connecting a fine capillary to a mercury reservoir and dipping the end of the capillary into the test solution. A mercury pool at the bottom of the cell served as both the counter and the reference electrodes in a single-compartment, two-electrode setup. In the classical configuration, the flow rate of mercury was determined by the hydrostatic pressure between the top of the mercury reservoir and the tip of the capillary, and by the dimensions of the capillary. Mercury drops were formed at the tip of the capillary, and were detached at almost equal intervals, when the weight of the drop just exceeded the surface tension holding it.

Mercury electrodes played a major role in the early development of electrochemistry, not only for electroanalytical purposes, but also for the understanding of the thermodynamics and kinetics of the metal/solution interphase, the potential dependence of the double-layer capacitance, and its relation to the dependence of the surface tension on potential. Electrosorption, the potential-dependent adsorption of ions and neutral molecules on electrode surfaces, was also studied initially on mercury electrodes. The unique features of mercury electrodes are that the metal can be purified by distillation, and the electrode is renewed periodically, each new drop presenting a highly reproducible, fresh, structureless, and clean surface. Moreover, mercury is a very poor catalyst for hydrogen evolution. Consequently, it acts as an (almost) ideally polarizable interface over a fairly wide range of potentials, particularly in the cathodic direction, allowing the study of many cathodic electrode reactions for which the standard potential is negative with respect to the standard hydrogen electrode, with little interference from decomposition of the solvent (i.e. of hydrogen evolution). Where metal deposition is concerned, one finds the added advantage that many metals are soluble in mercury, forming an amalgam. The potential for formation of the amalgam can be significantly shifted in the positive direction, compared to deposition on a solid electrode. For the alkali metals, this

shift can be as much as one volt, allowing the study of the rate of deposition of such metals on mercury, but not on solid metals.

Mercury was used in the past in large quantities in the chlor-alkali industry, for the production of chlorine and pure NaOH. However, its use in the industry was essentially stopped in many countries when its deadly impact on the environment was realized; even in research its use has been limited as far as possible. In modern polarography, pressure is applied by an inert gas, and the experiment is conducted using very small amounts of mercury.

For the classical dropping mercury electrode, the diffusion-limited current is given by the Ilkovic equation

$$I_d = 708\,\dot{m}^{2/3} t^{1/6} D^{1/2} c_b \tag{4.17}$$

where $\dot{m}$ is the flow rate of Hg (mg s^{-1}), t is the drop time (s), D is the diffusion coefficient (cm^2 s^{-1}), and c_b is the bulk concentration (mmol L^{-1}). The current calculated from Eq. (4.17) with these units is given in microampere. Under typical polarographic conditions, the value of I_d calculated from the Ilkovic equation is about 3 μA for $n = 1$, a drop time of 1 second, in a 1 mM solution. Note that this is a total current, I, not a current density, j, and the concentration is given in units of millimolar, in agreement with the accepted notation in the polarographic literature (in most other parts of electrode kinetics, the current density and the concentration are given in units of A cm^{-2} and mol cm^{-3}, respectively). The area of the drop, just before it falls, is about 0.01–0.04 cm^2, yielding current densities in the range of 0.8–3.0 mA cm^{-2} under the above conditions.

Classical dropping mercury electrode (DME) represents a rather complex system. As the volume of the drop grows, its surface moves toward the solution, while the diffusion layer is also growing. As a result, the thickness of the diffusion layer is less than that which would have been calculated for a stationary electrode. Therefore, the Ilkovic equation is an approximation, not an exact solution of the diffusion problem. The dependence of the diffusion-limited current on $t^{1/6}$ results from the combined effects of the increasing surface area (at a rate that is proportional to $t^{2/3}$) and the decreasing diffusion-limited current (that is proportional to $t^{-1/2}$).

For a reversible process, the dependence of the potential on the current is expressed by the equation

$$E = E_{1/2} + \frac{2.3RT}{nF} \log \frac{I_d - I}{I} \tag{4.18}$$

where $E_{1/2}$ is the polarographic half-wave potential, given by

$$E_{1/2} = E^0 + \frac{2.3RT}{nF} \log \left[\left(\frac{\gamma_{Ox}}{\gamma_{Red}} \right) \left(\frac{D_{Red}}{D_{Ox}} \right)^{1/2} \right] \tag{4.19}$$

where γ_i stands for the activity coefficient of the species indicated by the subscript. Since the logarithmic term on the right hand side of this equation is usually quite small, the measured value of $E_{1/2}$ is close to that of E^0, and it can serve to identify the species being reduced.

The current–potential relationship for irreversible polarographic waves is more complex than that for the reversible situation. Instead of Eq. (4.18), one can use the approximate expression

$$E = E_{1/2} + (0.916)b \log \frac{I_d - I}{I} \tag{4.20}$$

The half-wave potential in the above equation is given by

$$E_{1/2} = E^0 + b \log[1.32k_{s,h}(t/D)^{1/2}] \tag{4.21}$$

where $k_{s,h}$ is the heterogeneous rate constant at the standard potential, E^0, and b is the Tafel slope, to be discussed later in detail. The important point to note here is that the half-wave potential is a function of the heterogeneous rate constant and of the time of measurement; it is no longer determined by the standard potential alone.

In *classical polarography* (also referred to as DC polarography), the potential is changed linearly with time at a slow rate of a few *millivolts per second*, and the drops are allowed to fall freely. In modern instruments, the mercury drop is extruded in a few milliseconds to its final size, determined by the dimensions of the capillary and the gas pressure. The surface area is then maintained constant during measurement. This type of setup is referred to as *static dropping mercury electrode* (SDME).

In another method, called *normal-pulse polarography* (NPP), a commonly used technique, the current is sampled once on each drop, and the potential is applied for a short time, just before the drop is knocked off. For a pulse width in the range of 1–50 ms, the thickness of the diffusion layer is 10 μm or less, compared to a radius of curvature of 250–350 μm for typical mercury drops. The thin diffusion layer can increase the limiting current (and, hence, the sensitivity) by more than an order of magnitude, compared to DC polarography. Moreover, the system can be treated as semi-infinite linear diffusion, yielding an accurate equation for the diffusion-limited current, since the complications attributed earlier to the expansion of the electrode surface during the measurement have been eliminated. The limiting current under these conditions is given by

$$I_d = nFA_p c_b \left(\frac{D}{\pi t_p} \right)^{1/2} \tag{4.22}$$

where the A_p and t_p represent the area of the drop during the application of the pulse and the duration of the pulse, respectively.

Several other techniques related to polarography have been developed over the years, mainly to increase the sensitivity, but also to reduce the time of measurement and the amount of mercury used. The most common among them are *differential-pulse polarography* (DPP) and *square-wave polarography* (SWP). These methods of employing the SDME for analytical applications have been widely treated in the literature and will not be discussed here.

4.3 Rotating Electrodes

4.3.1 The Rotating Disk Electrode (RDE)

There are many ways to increase the rate of mass transport by stirring. Moving the electrode in solution turns out to be more efficient, as a rule, than moving the solution by gas bubbling or using a magnetic stirrer.

One of the best methods of obtaining efficient mass transport in a highly reproducible manner is by the use of the *rotating disk electrode* (RDE). The RDE consists of a cylindrical metal rod embedded in a larger cylindrical plastic holder (usually Teflon, due to its great chemical stability and inertness). The electrode is cut and polished flush with its holder, so that only the bottom end of the metal cylinder is exposed to the solution.

The configuration of an RDE is shown schematically in Figure 4.5a, compared to a rotating ring disk electrode (RRDE), which is shown in Figure 4.5b and discussed later (see Section 4.3.3). The most important feature of the RDE is that it acts as a so-called *uniformly accessible surface*, which, in simple language, means that the rate of mass transport to the surface is uniform. This property is by no means self-evident, considering that the linear velocity of points on the surface of the rotating disk increases with their distance from the center of rotation. The other important property of the rotating disk is that the flow of the solution at the surface is laminar up to rather high rotation rates.

In hydrodynamics, the transition from laminar to turbulent flow is characterized by a dimensionless parameter called the *Reynolds number*, *Re*:

$$Re = \frac{vl}{\upsilon} \tag{4.23}$$

where v is a characteristic velocity (cm s^{-1}), l is a characteristic length (cm), and υ is the kinematic viscosity, which is defined as $\upsilon \equiv \eta/\rho$, where η is the dynamic (absolute) viscosity (g cm^{-1} s^{-1}), and ρ is the density (g cm^{-3}). Note

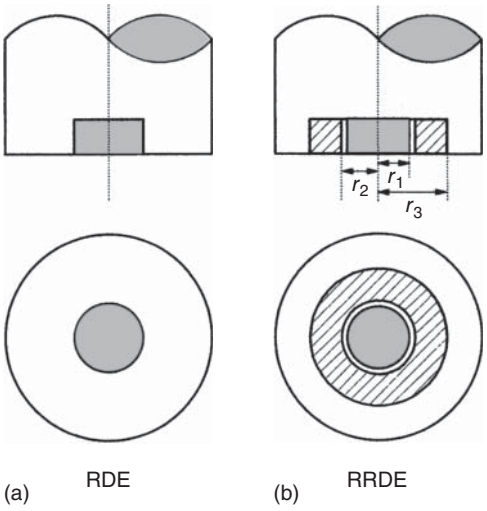


Figure 4.5 The rotating disk electrode (a) and the rotating ring disk electrode (b). The electrodes and insulating material are marked by filled and clear areas, respectively.

that the values of dynamic viscosity are often reported in the literature in units of centipoise (cP), where 1 cP = 0.01 poise (P), where 1 P = 1 g cm^{-1} s^{-1}. The kinematic viscosity has the dimensions of square centimeter per second (the same as the dimensions of the diffusion coefficient, D). The characteristic velocity and length must be defined for each geometry separately, and in each case there is a critical value of the Reynolds number at which transition from laminar to turbulent flow starts. For example, for flow in a tube: $l = r$. The corresponding critical Reynolds number is about 2×10^3. In dilute aqueous solutions at room temperature, $\upsilon \approx 10^{-2}$ cm^2 s^{-1}. Consequently, flow in a major pipeline is usually turbulent (for $r = 25$ cm and $\nu = 10$ cm s^{-1} one finds $Re = 2.5 \times 10^4$), whereas in small capillaries it is almost always laminar (for $r = 0.1$ cm and $\nu = 10$ cm s^{-1}, $Re = 1 \times 10^2$).

In the case of the rotating disk, the characteristic velocity is the linear velocity at the outer edge of the disk, given by $\nu = \omega r$, where ω is the angular velocity, expressed in radians per second. The characteristic length is taken as the radius of the disk and the critical Reynolds number is about 1×10^5. The transition from laminar to turbulent takes place at

$$Re = \omega r \frac{r}{\upsilon} = \frac{\omega r^2}{\upsilon} \leq 1 \times 10^5 \tag{4.24}$$

For a typical radius of 0.2 cm, laminar flow should be maintained up to $\omega \leq 2.5 \times 10^4$ rad s^{-1}, which corresponds to a rotation rate of about 2.5×10^5 rpm.

It should be noted that the critical Reynolds numbers represent, as a rule, the upper limits for laminar flow at ideally smooth surfaces. If the surface is rough, turbulence may set in at a lower Reynolds number. Most RDEs are operated at a maximum rotation rate of $N \leq 10^4$ rpm ($\omega \approx 1 \times 10^3$ rad s^{-1}), well within the range of laminar flow. The lower limit for the rotation rate is determined by the requirement that the limiting current density resulting from rotation should be large compared to that which would exist in a stagnant solution due to natural convection. In practice, this corresponds to a lower limit in the range of about 400 rpm, which can be extended to 100 rpm under carefully controlled experimental conditions.

One should be careful in applying the foregoing considerations to electrodes of widely different dimensions. For example, if one were to employ the RDE in an industrial cell using, say, an electrode area of 1000 cm^2, the critical Reynolds number would be reached at the rim of the electrode at a rotation rate of only 30 rpm, and turbulence could set in at even lower rotation rates. Such electrodes may still be of practical value in an industrial process, as long as one is aware that flow may become turbulent beyond a certain radius, and the surface will no longer be uniformly accessible. At the other extreme, the assumptions made in solving the hydrodynamic equations under most circumstances imply that one would be ill advised to use an RDE with a radius smaller than 0.05 cm.

Since flow is laminar, it is possible to calculate the rate of mass transport. The corresponding equation for the limiting current density, developed by Levich, is

$$j_\mathrm{L} = 0.62nFD^{2/3}\upsilon^{-1/6}\omega^{1/2}c_\mathrm{b} \tag{4.25}$$

The current density calculated from this equation is in ampere per square centimeter if the angular velocity is given in radians per second and the concentration

is given in mole per cubic centimeter. We can rewrite Eq. (4.25) in the more convenient form as

$$j_L = 0.201\, nF D^{2/3} \upsilon^{-1/6} N^{1/2} c_b \tag{4.26}$$

where the angular velocity is replaced by the rotation rate, N, in units of revolutions per minute. Comparing the last two equations with the equation for mass-transport-limited current density,

$$j_L = \frac{nFDc_b}{\delta} \tag{1.4}$$

we find, for the diffusion-layer thickness, δ, at the RDE, the expressions

$$\delta = 1.61 D^{1/3} \upsilon^{1/6} \omega^{-1/2} \tag{4.27}$$

with ω in radians per second and

$$\delta = 5.00 D^{1/3} \upsilon^{1/6} N^{-1/2} \tag{4.28}$$

for the rotation rate expressed in revolutions per minute.

For typical values of $D = 1 \times 10^{-5}$ cm^2 s^{-1} and $\upsilon = 1 \times 10^{-2}$ cm^2 s^{-1}, the diffusion-layer thickness turns out to be 5–50 μm for rotation rates of 10^2–10^4 rpm. This range should be compared to values of 50–150 μm obtained by stirring of the solution. For a diffusion-controlled process, a value of $d = 5$ μm is reached in a stagnant solution after about 8 ms (see Eq. (4.11)), showing the advantage of even moderately fast transients (say, in the range of 0.1–1.0 ms) in enhancing the rate of mass transport. On the other hand, measurement with an RDE is a steady-state measurement, which is often advantageous.

Considering Eq. (4.26), we note that j_L is a linear function of concentration. Hence, the RDE can be used as a tool in electroanalytical measurements. It has also been used extensively to determine the diffusion coefficients of different electroactive species in solution.

One may wonder why the movement of the RDE in a fixed plane has an effect on mass transport in a direction perpendicular to that plane. This comes about because the disk drags the solution nearest to it and imparts to it momentum in the tangential direction. As a result, solution is pushed out of the surface sideways (in a plane parallel to it) and is replaced by solution moving in from the bulk, in a direction perpendicular to the surface. The rotating surface acts, in effect, as a pump, pulling the liquid up toward it, as shown schematically in Figure 4.6.

So far, we have discussed the properties of the limiting current density at the RDE. What about activation or mixed control? For a purely activation-controlled process, the current should be independent of rotation rate, or should at least become independent of it beyond a certain rotation rate. Under conditions of mixed control, the activation-controlled and mass-transport-controlled current densities combine to yield the total current density as the sum of reciprocals, namely

$$\frac{1}{j} = \frac{1}{j_{ac}} + \frac{1}{j_L} \tag{1.2}$$

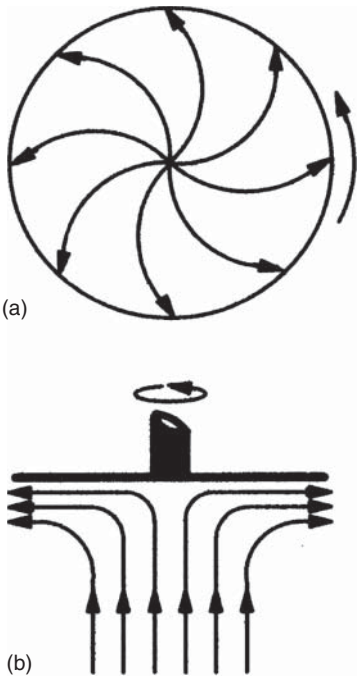


Figure 4.6 Motion of the liquid at the surface of an RDE. (a) Tangential motion in the plane of the electrode. (b) Perpendicular motion toward the electrode.

Since j_L is proportional to $\omega^{1/2}$, this can be rewritten as

$$\frac{1}{j} = \frac{1}{j_{ac}} + \frac{1}{B\omega^{1/2}} \tag{4.29}$$

Equation (4.29) is very useful for the study of electrode kinetics. It is clear that a plot of $1/j$ versus $1/\omega^{1/2}$ should yield a straight line having an intercept of $1/j_{ac}$. By repeating the experiment at different potentials, one can obtain the dependence of $1/j_{ac}$ on potential, as shown in Figure 4.7. Extrapolation of $1/j$ to $N^{-1/2} = 0$ is equivalent to measuring at a point where the rotation rate, and thus the limiting current density, is extrapolated to infinity. At this point, the current would evidently be controlled only by the rate of charge transfer.

The constant B, used in Eq. (4.29), appears to be independent of potential. This is indeed true in most cases, but a detailed analysis shows that it is a function of potential when the system is studied close to equilibrium. It is nevertheless reassuring to note that the value of j_{ac} obtained by extrapolating $1/j$ versus $N^{-1/2}$ in the foregoing manner is correct, irrespective of the dependence of the parameter B on potential.

As may be expected, several methods have been developed that make use of the basic idea of the RDE, but improve upon it or extend it, for the purpose of studying electrode kinetics. Some of these are discussed later. Some caution should be exercised, however, in the analysis of such data, because of the nonuniformity of current distribution. Thus, although the RDE represents a *uniformly accessible surface*, implying that the mass-transport-limited current density is uniform, the same cannot be said for the activation-controlled current density. The configuration of a metal electrode and its insulating holder in the same plane necessarily

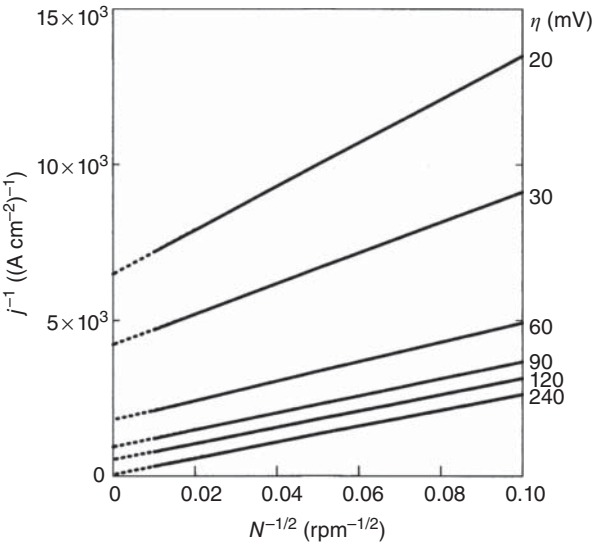


Figure 4.7 Analysis of the results obtained with an RDE, in the region of mixed control. The activation-controlled current density is calculated from the intercept of the lines at infinite rotation rate, where $N^{-1/2} \rightarrow 0$.

leads to large edge effects (see Figures 3.6 and 3.7), if the rate of mass transport is high, so that its effect on the measured current density is small. The resulting error can be reduced, but not entirely eliminated, by placing a ring around the electrode, separated from it electrically and held at the same potential.

The best way to avoid the problem of nonuniform current distribution on the electrode surface is by operating at lower current densities, which increases the value of R_F, or by increasing the conductivity in solution (to decrease R_S), or both.

4.3.2 The Rotating Cone Electrode (RConeE)

The usefulness of the RDE stimulated the development of several other rotating configurations that are worth mentioning. Some of these have evolved in response to specific experimental needs, whereas others serve as possible extensions of the technique.

Often, bubbles are generated at the electrode surface (e.g. when the reaction being studied is a gas evolution reaction) or reach it from the bulk of the solution where gas is used to de-aerate it. A bubble trapped at the surface of the RDE is driven toward its center by centrifugal force. This phenomenon can cause major errors in measurement because: (i) part of the surface of the electrode is blocked, and (ii) the hydrodynamic flow in the vicinity of the electrode may be greatly distorted.

A solution to this problem is to use a *rotating cone electrode* (RConeE). It turns out that this electrode is similar to an RDE in that it acts as a uniformly accessible surface, producing a uniform rate of mass transport to all parts of the electrode

surface. For a cone having an opening angle of θ, the current density is given by

$$j_{\mathrm{L,cone}} = j_{\mathrm{L,disc}} \sqrt{\sin(\theta/2)} \tag{4.30}$$

When a bubble forms at the rotating cone, it is driven to its center, namely to the tip of the cone. However, it is in a rather precarious state there, and if it is not dislodged spontaneously by some minor instability in rotation, it can be easily removed.

4.3.3 The Rotating Ring Disk Electrode (RRDE)

The rotating ring disk electrode (RRDE) is perhaps the most useful extension of the idea of the rotating disk. The great advantage of the RRDE is that it can be used to analyze short-lived intermediates in a steady-state measurement. The RRDE is constructed in such a way that both electrodes are in the same plane, and in close proximity to each other. The dimensions of the electrode are defined by three radii: the radius of the disk (r_1), the inner radius of the ring (r_2), and its outer radius (r_3), as shown in Figure 4.5b.

The ring current has the same form as the disk current, namely, it is also proportional to the bulk concentration of the electroactive species and to the square root of the rotation rate:[2]

$$I_{\mathrm{R}} = 0.62nF\pi(r_3^3 - r_2^3)^{2/3} D^{2/3} \upsilon^{-1/6} \omega^{1/2} c_{\mathrm{b}} \tag{4.31}$$

In this equation, I_{R} is in units of Ampere. Since both electrodes are part of the same flow regime (evidently, the convective flow of the liquid cannot distinguish between the disk electrode, the insulator and the ring), the equations for the limiting current at both the disk and the ring electrodes can be solved simultaneously.

In order to operate an RRDE, one needs a four-electrode potentiostat, often referred to also as a "*bi-potentiostat*." Now, we began this book with simple two-electrode cells and proceeded to the three-electrode configuration. How does the fourth electrode fit into this scheme? The best way to understand this is to view both the ring and the disk as working electrodes, with common counter and reference electrodes. The important point is that the four-electrode potentiostat controls the potentials of the disk and of the ring with respect to the reference electrode *independently* and measures the current going through each of them separately. We could, for example, hold the potential of the ring at a fixed value and vary the potential of the disk, allowing an analysis of the products formed at the disk by studying the ring current as a function of the potential applied to the disk, or vice versa.

The RRDE is calibrated by determining its collection efficiency, N. This parameter represents the fraction of a species formed at the disk that reaches the ring and reacts there. To do this, consider a simple oxidation reaction, say, Fe^{2+}/Fe^{3+}. The disk potential is set at a positive value where the oxidation process occurs at some rate I, while the ring potential is set at a negative value at which the oxidized species is reduced at its limiting current. If there is initially no Fe^{3+} in solution,

2 Note that for the rotating ring and rotating ring disk electrodes the common practice is to use the total currents, I, rather than the current densities, j.

the ring current results only from Fe^{3+} ions produced at the disk and transported to the ring by convection (there is always some diffusion, but its rate is negligible compared to the rate of convection in an RRDE experiment). The ratio of ring current to disk current observed under these conditions is defined as the collection efficiency, N

$$N = I_R/I_D = f(r_1, r_2, r_3) \tag{4.32}$$

where I_R and I_D represent the *total* currents measured for the ring and the disk, *not* the corresponding current densities. The important thing to notice in this equation is that the collection efficiency is not a function of the rotation rate; rather, it depends exclusively on the dimensions of the electrodes. The function $f(r_1, r_2, r_3)$ is somewhat complicated, but its values have been tabulated, and it can be obtained given the values of the above three radii. In practice, it is always preferable to evaluate N experimentally, since it depends strongly on the width of the gap, $r_3 - r_2$, which cannot be measured very accurately. Typical values of the collection efficiency for commercial RRDEs are in the range of 0.1–0.3. Decreasing the insulating gap between the disk and the ring increases the value of the collection efficiency. Decreasing the width of the ring ($r_3 - r_2$) leads to a lower collection efficiency, but improves the accuracy of measurement of the lifetime of unstable intermediates formed at the disk.

If an *unstable intermediate or product* is formed at the disk, only a fraction of it will reach the ring, and the ratio I_R/I_D will be smaller than N. In this case, the extent to which the collection efficiency has decreased is a function of the rate of rotation. The dependence of the collection efficiency on the rotation rate can be used to evaluate the lifetime (or rate of decomposition) of the unstable intermediate. As in any kinetic measurement of this type, one attempts to design the system for the fastest possible transition time from disk to ring, to allow detection of short-lived intermediates. The gap in commercial RRDEs is on the order of 0.01 cm, but electrodes with substantially narrower gaps have been built. One might be tempted to use modern techniques of microelectronics to construct an RRDE with a very small gap, say 0.1 μm. A closer examination of the hydrodynamics involved reveals, however, that this might not work and, in fact, there is little or no advantage in reducing the gap below about 5–10 μm.

The RRDE can also be very useful in other studies, where the species formed at the disk are stable. For example, the corrosion of a metal is always associated with a cathodic reaction, which in de-aerated solutions is hydrogen evolution, while in the presence of air it might be the reduction of molecular oxygen. The rate of corrosion of a metal (used as the disk) at open circuit can therefore be followed indirectly by measuring the amount of molecular hydrogen reaching the ring electrode or by the decline in the amount of molecular oxygen reaching it. The RRDE can also be used to study complex reactions in which two processes occur in parallel. For example, the electrodeposition of zinc is accompanied by hydrogen evolution. The partial current for hydrogen evolution at the disk can then be determined by measuring the amount of molecular hydrogen reaching the ring. Similarly, when a $Ni(OH)_2$/NiOOH electrode (used in a Ni/Cd battery) is charged, oxygen evolution is known to start before the electrode has been fully charged. The partial current for charging can be determined by monitoring the

amount of oxygen reaching the ring electrode, and calculating the corresponding current, corrected for the collection efficiency, from the total observed disk current.

Further Reading

1 Bard, A.J., Inzelt, G., and Scholz, F. (ed.) (2012). *Electrochemical Dictionary*. Berlin Heidelberg: Springer-Verlag.
2 Calvo, E.J. (ed.) (2007). *Encyclopedia of Electrochemistry*, Interfacial Kinetics and Mass Transport, vol. 2. Weinheim: Wiley-VCH.
3 Bard, A.J. and Faulkner, L.R. (2001). *Electrochemical Methods: Fundamentals and Applications*, 2nd edn. New York, USA: Wiley.
4 Gileadi, E. (1993). *Electrode Kinetics – for Chemists, Chemical Engineers and Materials Scientists*. Wiley-VCH.
5 Prentice, G. (1991). *Electrochemical Engineering Principles*. New Jersey, USA: Prentice-Hall.
6 Gileadi, E., Kirowa-Eisner, E., and Penciner, J. (1975). *Interfacial Electrochemistry – An Experimental Approach*. Addison-Wesley.
7 Newman, J.S. (1973). *Electrochemical Systems*. New Jersey, USA: Prentice-Hall.

5

Single-Step Electrode Reactions

5.1 The Overpotential, η

5.1.1 Definition and Physical Meaning of Overpotential

The standard potential for a given electrode reaction is related to the difference in the Gibbs free energy between reactants and products. Its dependence on the concentrations of reactants and products in solution is expressed by the Nernst equation, which has the general form

$$E_{\text{rev}} = E^0 + \frac{2.3RT}{nF} \log \frac{c_{\text{Ox}}}{c_{\text{Red}}} \tag{5.1}$$

where the activity coefficient is usually taken as unity, for convenience. When at open circuit, a nonpolarizable electrode assumes its reversible potential, whereas a polarizable electrode may deviate from it significantly. In either case, the overpotential, η, is defined as the difference between the actual potential measured (or applied) and the reversible potential:

$$\eta \equiv E_{\text{app}} - E_{\text{rev}} \tag{5.2}$$

where E_{app} is the applied potential. This definition of overpotential is phenomenological and is always valid, irrespective of the reasons for the deviation of the potential from its reversible value. The overpotential is always defined with respect to a specific reaction, for which the reversible potential is known. When more than one reaction can occur simultaneously on the same electrode, there is a different overpotential with respect to each reaction, for any value of the measured potential. This situation is encountered, *by definition*, in the course of alloy deposition, in which at least two metals are being deposited simultaneously in parallel. Moreover, in many cases, additives such as ammonia or citric acid are added to improve the quality of the plating process. When NH_4OH is used in the plating of nickel, for example, several complexes of the type $[Ni(NH_3)_n]^{2+}$ may exist, where n can assume values of 2, 3, 4, and 6. Each of these complexes will have a different reversible potential. In such situations, analysis of the mechanism based on measuring the Tafel slope is questionable.

The overpotential is a measure of the distance of a reaction, on the Gibbs energy scale, from its equilibrium state. Multiplied by the charge consumed per mole of reactants, nF, it is equal to the affinity, A, of a reaction, which is a measure of the

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

thermodynamic driving force that exists to make the reaction occur:

$$A = nF|\eta| \tag{5.3}$$

Now, we have already seen in Chapter 4 that the electrochemical Gibbs free energy of activation is linearly related to the applied potential, giving us a powerful tool to control the rate of electrode reactions over many orders of magnitude. At the other extreme, we can also use the potential to probe the reaction under conditions close to equilibrium, by applying small values of the overpotential in both directions around zero, and measuring the resulting current densities.

It should be recalled that equilibrium does not imply total freezing of the reaction. It is characterized by a dynamic situation, in which the reaction proceeds in both directions. Equilibrium is reached when the forward and backward rates are equal and the *net rate*, which one would observe, say, as a change of concentration in a chemical reaction or a flow of current in the external circuit in an electrochemical reaction, is zero.

Consider the following simple chemical reaction at equilibrium, with reactants and products all in the gas phase:

$$\mathrm{I_2 + H_2 \rightleftharpoons 2HI} \tag{5.4}$$

By definition, the affinity $A = 0$ and the net reaction rate is $\nu = 0$. We can disturb the equilibrium slightly by adding a small amount of one of the reactants or by changing the temperature. The rate at which the reaction will proceed toward its new equilibrium is given by

$$\frac{\nu}{\nu_{\mathrm{eq}}} = \frac{A}{\mathrm{RT}} \tag{5.5}$$

where ν_{eq} is the exchange rate, namely, the rate at which the reaction proceeds back and forth at equilibrium.

For transfer of a unit charge (i.e. an electron or an ion) we can similarly write

$$\frac{j}{j_0} = \frac{\eta F}{\mathrm{RT}} \tag{5.6}$$

In this equation, j_0 is called the *exchange current density*, representing the current density in both anodic and cathodic directions when the system is at equilibrium.

Equations (5.5) and (5.6) are special cases of a general rule, according to which, whenever a system is perturbed to a small extent, the response is a linear function of the magnitude of the perturbation. But, how small is "small" in the present context? This can be defined in dimensionless form, comparing affinity to the average thermal energy, or the rate observed to the exchange rate. For Eq. (5.5) the perturbation is small if $A/RT \ll 1$. Likewise, for the case of charge transfer, the requirement is $|\eta|F/RT \ll 1$. Clearly, a small perturbation also leads to $\nu/\nu_{\mathrm{eq}} \ll 1$ and $j/j_0 \ll 1$. However, this latter criterion is mainly relevant to electrode kinetics, since only there can one control the reaction rate (in a galvanostatic measurement) and observe the resulting affinity (i.e. overpotential).

To obtain a better feel for the quantities involved, we note that RT, the average thermal energy, at 298 K is equal to 25.7 meV (2.48 kJ mol^{-1}). Thus, for $\eta = 5$ mV we have

$$\frac{|\eta|F}{RT} \approx 0.2 \tag{5.7}$$

The value of the exchange current density, j_0, can readily be obtained from the slope of a plot of j versus η, using Eq. (5.6).

5.1.2 Types of Overpotential

In the previous section we viewed the overpotential as the stimulus (or perturbation), which causes the reaction to proceed in the desired direction, at a certain rate, represented by j. The ratio between the perturbation and the response is a measure of the effective resistance of the system, called the faradaic resistance, R_F (also referred to as the charge-transfer resistance, R_{ct}). In the example given earlier, which is applicable to small perturbations, the ratio η/j is constant, since the current is a linear function of the overpotential. However, in general, R_F is a function of the potential, as seen in Figure 1.1. The units of R_F are the product of the resistance and the area, namely Ω cm^2, since it is defined as the ratio between the overpotential and the current density. Multiplied by the current density, in units of A cm^{-2}, it yields the overpotential

$$jR_F = \eta_F \tag{5.8}$$

In general, it is better to define the faradaic resistance in differential form as

$$R_F \equiv \left(\frac{\partial \eta}{\partial j}\right)_c \tag{5.9}$$

bearing in mind that R_F is a function of η, as seen in Figure 1.1.

The overpotential observed experimentally can be the result of unrelated physical phenomena, each of which can influence the reaction rate. Here, we shall discuss four components.

Activation overpotential, η_{ac} (also called the charge-transfer overpotential, η_{ct}) is the kinetically significant quantity acting on the energy of activation, lowering it in one direction and increasing it in the opposite direction; η_{ac} is the change in potential across the interphase generated by an applied current density.

$$\eta_{ac} = {}^{M}\Delta^{S}\phi - {}^{M}\Delta^{S}\phi_{rev} = \delta\left({}^{M}\Delta^{S}\phi\right) \tag{5.10}$$

In proper electrode kinetic studies one needs to isolate this type of overpotential. The mechanistic information is then based on determining the dependence of η_{ac} on current density over a wide range of experimental conditions.

Concentration overpotential, η_{conc}, is the second important type of overpotential caused by mass-transport limitation. When a reaction takes place, the concentration of the reacting species at the surface is different from its value in the bulk (see Figure 4.1). This manifests itself by an added overpotential related to the limiting current density, by the equation

$$\eta_{conc} = \frac{2.3RT}{nF} \log\left(1 - \frac{j}{j_L}\right) \tag{5.11}$$

Resistance (Ohmic) overpotential, η_{jR_S}, is the third type of overpotential that results from the residual potential drop jR_S in the solution between the working electrode and the tip of the Luggin capillary connected to the reference electrode. In the research laboratory, this type of overpotential can be largely eliminated by proper positioning of the Luggin capillary, combined with electronic compensation for most of the remaining resistance, in real time. It does, however, have a major effect in industrial electrochemistry, with respect to its influence on the performance of batteries and on the energy consumption of electrolytic processes.

This energy is transformed to heat, which must be removed by proper heat exchangers in order to maintain the desired temperature of operation. The immense magnitude of this effect may be appreciated when it is realized that the amount of electricity transformed into thermal energy in the chlorine and aluminum industries combined is about 2–3% of the total electric power production in the United States alone.

The resistance overpotential increases as either the resistivity of the solution or the applied current increases. This overpotential can also be very important in localized crevice and pitting corrosion phenomena, in the application of cathodic and anodic protection, or when nonconductive surface films, such as oxides and paints, are present.

Crystallization overpotential (η_{cryst}) is observed when a metal is deposited on a foreign metal substrate at a potential negative to the reversible potential for the system studied. This is the equivalent of suppercooling in phase transformation. For example, water can be cooled down to −10 °C without freezing; it is maintained in a metastable phase. One or more of the above phenomena can add to the total overpotential observed as

$$\eta = \eta_{ac} + \eta_{conc} + \eta_{jR_S} + \eta_{cryst} \tag{5.12}$$

This equation is conceptually correct, but it should be considered that the above types of overpotential may interact, and an attempt to obtain η by calculating the four contributions separately and adding them might lead to some errors.

5.2 Fundamental Equations of Electrode Kinetics

5.2.1 The Empirical Tafel Equation

The *Tafel equation* was first published by Julius Tafel in 1905. In its simple form, it can be written as

$$\eta = a - b \log j \tag{5.13}$$

where a and b are empirical constants. This was, at the time, an empirical equation based on measurement of the hydrogen evolution reaction (HER) on Hg and Pb cathodes. The parameter b is evidently the *Tafel slope*, given by the equation

$$b = \left(\frac{\partial \eta}{\partial \log j}\right)_{T,P,c} \tag{5.14}$$

Since then, the above two equations have probably been the most frequently used equations in the studies of electrode kinetics.

The numerical value of b observed by Tafel was about 0.12 V per decade of current density. This is consistent with our previous discussion, in which we showed that the rate of an electrochemical reaction can be increased by many orders of magnitude by applying a relatively small overpotential to the electrode.[1] The conditions under which the Tafel equation can be observed, such as a high overpotential, no mass-transport limitation, uniform current distribution, and no interference by film formation, all happen to exist in the particular experimental systems chosen by Professor Tafel.

For many years, the Tafel equation was viewed as an empirical equation. A theoretical interpretation was proposed only after Eyring, Polanyi, and Horiuti developed the Transition-State Theory for chemical kinetics, in the early 1930s. Since the Tafel equation is one of the most important fundamental equations of electrode kinetics, we shall derive it first for a single-step process, and then extend the treatment for multiple consecutive steps. Before we do that, however, we shall review very briefly the derivation of the equations of the Transition-State Theory of chemical kinetics.

5.2.2 The Transition-State Theory

The *Transition-State Theory*, which is also known as the *Absolute Rate Theory* or the *Activated Complex Theory*, is an important general theory that has been adapted for electrode kinetics.

Consider the simple isotope-exchange reaction, in which a hydrogen atom reacts with a deuterium molecule, D_2, to form a molecule of HD and a free deuterium atom:

$$D—D + H \rightarrow D—H + D \tag{5.15}$$

One could imagine this reaction occurring by D_2 first splitting into two atoms, followed by one of the atoms combining with the hydrogen atom. This is, however, a highly unlikely course of events, because a great deal of energy is required to break the D—D bond. A probable route is the formation of an intermediate, such as an unstable D—D—H species, in which both bonds are of about equal strength, and the dissociation of this intermediate takes place to either the original reactants or to the products. The standard Gibbs free energy needed to form the unstable intermediate, which is called *the activated complex*, is the standard Gibbs free energy of activation for the reaction.

We shall now write the reaction in more general form as

$$A \rightarrow X^{\ddagger} \rightarrow B \tag{5.16}$$

in which A and B are the reactant and product molecules, respectively, and $X^{\ddagger}$ represents the activated complex. Two main assumptions are made in the framework of the Transition-State Theory: (i) the reaction rate is proportional to the

1 The December 2005 issue of the journal *Corrosion Science* was dedicated to the 100th anniversary of the publication of the Tafel equation. It contains mostly review articles, which can be useful for a deeper understanding of electrode kinetics.

concentration of the activated complex, and (ii) the activated complex is at equilibrium with the reactant. With these assumptions one can write the reaction rate as follows:

$$v = k_f c_A = \frac{kT}{h} c_{\ddagger} \tag{5.17}$$

For equilibrium between the reactant and the activated complex we can write

$$\frac{c_{\ddagger}}{c_A} = K^{\ddagger} = \exp\left(-\frac{\Delta G^{0\ddagger}}{RT}\right) \tag{5.18}$$

In Eq. (5.17), k_f, k, and h are the forward rate constant, the Boltzmann constant, and the Planck constant, respectively. Equation (5.18) shows the standard Gibbs free energy of activation, $\Delta G^{0\ddagger}$. The latter equals, by definition, the standard Gibbs free energy of the reaction in which the activated complex is formed from the reactants. The term kT/h in Eq. (5.17) has units of frequency (s^{-1}) and follows from an added assumption, namely that the critical bond in the activated complex, which must be severed in order to form the product, is very weak and behaves "classically" (in the sense that the energy levels are so close to each other that there appears to be a continuum). Substituting $c_{\ddagger}$ from Eq. (5.18) into Eq. (5.17) we obtain

$$v = k_f c_A = \frac{kT}{h} c_A \exp\left(-\frac{\Delta G^{0\ddagger}}{RT}\right) \tag{5.19}$$

or

$$k_f = \frac{kT}{h} \exp\left(-\frac{\Delta G^{0\ddagger}}{RT}\right) \tag{5.20}$$

This equation shows the chemical rate constant as a function of the Gibbs free energy of activation. In order to use this equation in electrode kinetics, it is necessary to relate the standard Gibbs free energy of activation to the potential difference, $\Delta\phi$, across the interface.

5.2.3 The Equation for a Single-Step Electrode Reaction

The relationship between the Gibbs free energy and potential was discussed in Section 4.1.2. For the standard electrochemical Gibbs free energy of a reaction, we write Eq. (4.6), and for the standard electrochemical Gibbs free energy of activation – Eq. (4.7):

$$\Delta\overline{G}^{0} = \Delta G^{0} \mp zF\Delta\phi \tag{4.6}$$

$$\Delta\overline{G}^{0\ddagger} = \Delta G^{0\ddagger} \pm \beta F\Delta\phi \tag{4.7}$$

The symmetry factor β in Eq. (4.7) is discussed in Section 5.3. The positive and negative signs in Eq. (4.7) are applicable to cathodic and anodic reactions, respectively. When Eq. (4.7) is combined with the rate equation, we get, for the anodic process,

$$j = Fv = F\frac{kT}{h} c_A \left(\frac{\Delta G^{0\ddagger}}{RT}\right) \exp\left(\frac{\beta F\Delta\phi}{RT}\right) \tag{5.21}$$

Equation (5.21) can also be written as

$$j = Fk_{\mathrm{h}}c_{\mathrm{A}} \exp\left(\frac{\beta F}{RT}\eta\right) \tag{5.22}$$

in which k_{h} is the heterogeneous electrochemical rate constant at the reversible potential. Equation (5.21) is an expression for the current density (which, we recall, is the rate of the reaction in electrical units of current density) in terms of the concentration and the overpotential, both of which can be measured and controlled experimentally. It should be noted that in deriving Eq. (5.21) we have tacitly assumed that the concentration at the surface is independent of the current density; in other words, we have ignored mass-transport limitation. Strictly speaking, we should therefore replace η by η_{ac} in these equations. For the sake of simplicity we shall leave it as it is and refer specifically to mass-transport limitation separately.

Equation (5.22) represents the rate of an anodic (oxidation) reaction, for which the overpotential is, by definition, positive. For a cathodic reduction we write a similar equation, with a negative sign in the exponent. Thus, in both cases the current density increases exponentially with increasing absolute value of the overpotential, $|\eta|$.

The complete expression for the current density is obtained as the difference between the anodic and cathodic current densities.

$$j = j_{\mathrm{a}} - j_{\mathrm{c}} = Fk_{\mathrm{h,f}}c_{\mathrm{A}} \exp\left(\frac{\beta F}{RT}\eta\right) - Fk_{\mathrm{h,b}}c_{\mathrm{A}} \exp\left[-\frac{(1-\beta)F}{RT}\eta\right] \tag{5.23}$$

This is the current measured in the external circuit.

We noted earlier that the symmetry factor β must have a value between zero and unity. Also, we have seen in Eq. (4.9) that β represents the ratio of the effect of potential on the electrochemical Gibbs free energy of activation to its effect on the electrochemical Gibbs energy of the overall reaction. If this fraction is β in the forward (anodic) direction, it must be $(1 - \beta)$ in the backward (cathodic) direction.

When the reaction is at equilibrium, there is no current flowing in the external circuit. Equation (5.23) can be reduced to

$$Fk_{\mathrm{h,f}}c_{\mathrm{A}} = Fk_{\mathrm{h,b}}c_{\mathrm{B}} \equiv j_0 \tag{5.24}$$

Combining Eqs. (5.22) and (5.23) we obtain

$$\frac{j}{j_0} = \exp\left(\frac{\beta F}{RT}\eta\right) - \exp\left[-\frac{(1-\beta)F}{RT}\eta\right] \tag{5.25}$$

The physical meaning of $j_{0,}$ which is called the *exchange current density*, should be clear from the definition given in Eq. (5.25). Equation (5.25) is nowadays known as the Butler–Volmer equation (although it was originally derived in 1930 by Erdey-Grúz and Volmer). It is valid only for activation-controlled reactions.

Equation (5.25) is an expression for the current density in terms of the concentration and the overpotential, both of which can be measured and controlled experimentally. Strictly speaking, we should therefore replace η by η_{ac}

in Eqs. (5.24) and (5.25). For the sake of simplicity we shall leave it as it is, and refer specifically to mass-transport limitation separately. For both the anodic (oxidation) reaction and the cathodic (reduction) reaction, the current density increases exponentially with increasing absolute value of the overpotential, $|\eta|$.

The *exchange current density*, j_0, represents the rate at which the electrochemical reaction proceeds back and forth at equilibrium, when the net reaction rate, observed as a current flowing through the external circuit, is zero. It is similar to the exchange rate discussed earlier in connection with Eq. (5.6). We also note that j_0 is the heterogeneous electrochemical rate constant at $\eta = 0$, multiplied by the concentration. One cannot overemphasize the fact that, while the measured current is zero, the reaction is not "frozen," and can occur at a high rate. The exchange current density can be much larger than the current measured, which may be a small difference between large anodic and cathodic currents.

5.2.4 Limiting Cases of the General Equation

The Butler–Volmer equation has two limiting cases that are used in measuring the Tafel slope and the polarization resistance. If the overpotential is small, a linear approximation may be applied, as shown in Eq. (5.6):

$$\frac{j}{j_0} = \frac{\eta F}{RT} \tag{5.6}$$

In Figure 5.1 we show this linear region up to about ±20 mV. Three different values of the symmetry factor are shown. For $\beta = 0.50$, the line is strictly

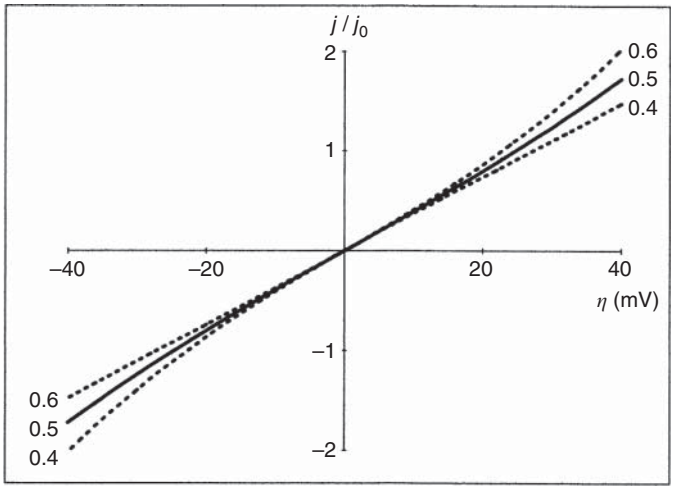


Figure 5.1 Plots of current density versus overpotential, calculated using Eq. (5.6) for a single-step charge-transfer process at low overpotentials. $\beta = 0.4$, 0.5, or 0.6. Note that linearity is maintained longer for $\beta = 0.50$ than for other values of β.

symmetrical. For two other values of β, the linear region is somewhat shorter, and the line is no longer symmetrical.

It is interesting to note that the symmetry factor disappeared from Eq. (5.6) in the process of linearization. Thus, the rate of a reaction close to equilibrium does not depend on the detailed shape of the energy barrier for activation (which determines the value of β). It does, however, depend on the magnitude of the energy of activation, which manifests itself in the value of j_0.

What is the range of the linearity represented by Eq. (5.6)? There is no single answer to this question, since it depends on the level of accuracy desired. Figure 5.1 shows that it is valid for about ±20 mV if the symmetry factor is 0.50, and up to about ±15 mV for values of $\beta = 0.40$ and 0.60.

We turn now to the high-overpotential region, as shown in Figure 5.2. One of the two exponential terms in Eq. (5.25) becomes negligible. For an anodic process, we have the equation

$$\frac{j}{j_0} = \exp\left(\frac{\beta F}{RT}\eta\right) \tag{5.26}$$

Written in logarithmic form, this equation becomes

$$2.3 \log \frac{j}{j_0} = \left(\frac{\beta F}{RT}\right)\eta = \frac{\eta}{b} \tag{5.27}$$

where the Tafel slope b is given by

$$b = \frac{2.3RT}{\beta F} \tag{5.28}$$

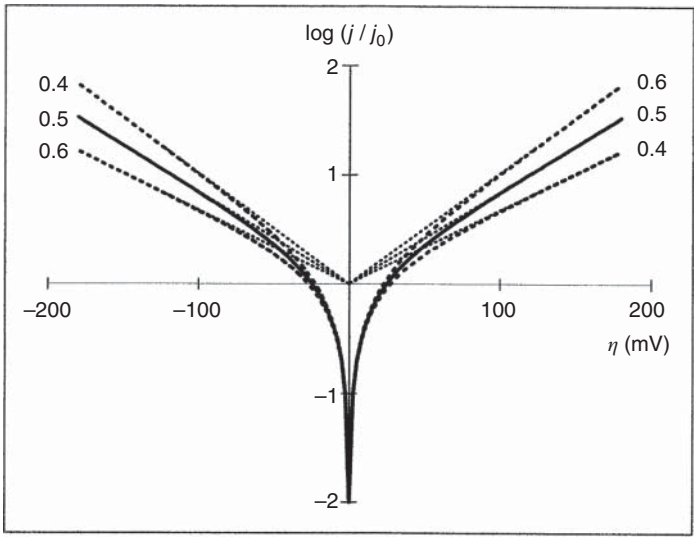


Figure 5.2 Semi-logarithmic plots of j/η, calculated from Eq. (5.27) for three values of β. Note that the lines are symmetrical only for $\beta = 0.50$.

Figure 5.2 shows the current–potential relationship based on Eq. (5.27), for a much larger range of potential. Note that, here, the current density is plotted on a logarithmic scale, yielding a straight line at high overpotential and curving toward the low values of η. One of the two exponential terms in Eq. (5.25) becomes negligible with respect to the other. For a large anodic (positive) overpotential one obtains

$$\frac{j_a}{j_0} = \exp\left(\frac{\beta F}{RT}\eta\right) \tag{5.29}$$

Written in logarithmic form, this leads to the equation

$$\frac{j_a}{j_0} = \frac{\beta F}{2.3\,RT}\eta = \frac{\eta}{b} \tag{5.30}$$

where the *Tafel slope b* (in units of V per decade of current density) is given as

$$b = \frac{2.3RT}{\beta F} \tag{5.31}$$

Equation (5.30) can also be written in the form

$$\eta = b\log\left(\frac{j}{j_0}\right) \tag{5.32}$$

which is the same as Eq. (5.13), just choosing the positive sign because here we refer to an anodic reaction.

It should be noted that, as in Figure 5.1, the symmetry of the lines is maintained only when $\beta = 0.50$. In Figure 5.2 we showed the j/η relationship in both anodic and cathodic directions on a semi-logarithmic scale.

The Tafel region, where η is a linear function of $2.3\log j$, occurs at high values of η. As for the low overpotential region, this evokes the question, how high is high? The rule of the thumb is to assume that this applies when $\eta/b \geq 1$. In Figure 5.2, the value of b was taken as $0.12\,\text{V decade}^{-1}$, which is the value obtained in the original work of Tafel.

A close look at Figure 5.2 reveals several interesting points:

(i) The Tafel plot is linear only at high overpotential, where $\eta/b \geq 1$.
(ii) The extrapolated anodic and cathodic lines intersect at $\eta = 0$, irrespective of the value of β. This property could serve to determine the reversible potential, but it rarely does, because in most cases it is difficult to determine the current–potential relationship both at high anodic and at high cathodic overpotentials. However, the reversible potential for the reaction studied is usually known, so that the value of j_0 can be determined experimentally by extrapolating the Tafel line to E_{rev}.
(iii) The Tafel plot is presented in terms of the current density, while the quantity determined experimentally is the total current. An uncertainty in the real surface area of the electrode (which often exists, as we shall see) causes a comparable uncertainty in the calculated value of j_0, but does not affect the Tafel slope. Similarly, if we do not know the reversible potential, we cannot determine j_0, but the Tafel slope can still be measured.

5.3 The Symmetry Factor, β, in Electrode Kinetics

5.3.1 The Definition of β

The symmetry factor has already been defined in Chapter 4 in terms of the ratio between the effect of potential on the electrochemical Gibbs free energy of activation and its effect on the electrochemical Gibbs free energy of the reaction

$$\beta \equiv \frac{(\delta\Delta\overline{G}^{0\ddagger}/\delta\Delta\phi)}{(\delta\Delta\overline{G}^{0}/\delta\Delta\phi)} = \frac{\delta\Delta\overline{G}^{0\ddagger}}{\delta\Delta\overline{G}^{0}} \tag{4.9}$$

It is appropriate to justify this definition here, and to compare it to some other ways in which this quantity has been defined in the literature. To do this, consider a simple reaction of the type

$$Ag^{+}_{soln} + e^{-}_{M} \rightarrow Ag^{0}_{M} \tag{5.33}$$

To be specific, we must note the positions of the various species before and after the reaction has taken place. The Ag^{+}_{soln} ion is on the solution side of the interphase, at a potential ϕ^{S} (which is also at the outer Helmholtz plane, which is also the distance of closest approach of the reactant to the surface). The electron is in the metal, at a potential ϕ^{M}. Following charge transfer, the silver atom will become part of the electrode and will be at the same potential ϕ^{M}, but since it is a neutral species, it is not affected by the potential, and we need not concern ourselves with it further. For the reaction to occur, a unit charge must cross the interphase, a process that requires electrical energy amounting to $F({}^{M}\Delta^{S}\phi)$. It does not matter, for the purpose of this argument, whether the reaction occurs by way of the positive ion crossing the interphase (from ϕ^{S} to ϕ^{M}) to be neutralized by the electron in the metal, or whether the electron jumps across, to neutralize the ion on the solution side of the interphase, or whether the particles meet somewhere in between. The electrochemical Gibbs free energy is *a function of state* defined thermodynamically only by its position. Thus, changes in it depend only on the initial and final states, not on the route the system has taken to go from one to the other.

We do not know the nature of the activated complex in such a process, but fortunately we do not need to know, for the present purpose. All we have to assume is that it is an intermediate state, somewhere between the initial and the final states, which by definition it is. Thus, we can say that the activated complex occurs at a point where the potential has an intermediate value, $\phi^{\ddagger}$, between ϕ^{M} and ϕ^{S}.

Since the formation of the activated complex is treated as an electrochemical reaction, we can write as follows for an anodic reaction:

$$\Delta\overline{G}^{0\ddagger} = \Delta G^{0\ddagger} - F\Delta\phi^{\ddagger} \tag{5.34}$$

and for the overall reaction we have, as before

$$\Delta\overline{G}^{0} = \Delta G^{0} \mp zF\Delta\phi \tag{4.6}$$

Hence, the symmetry factor, β, can be written as the ratio

$$\beta = \frac{\delta\Delta\phi^{\ddagger}}{\delta\Delta\phi} \tag{5.35}$$

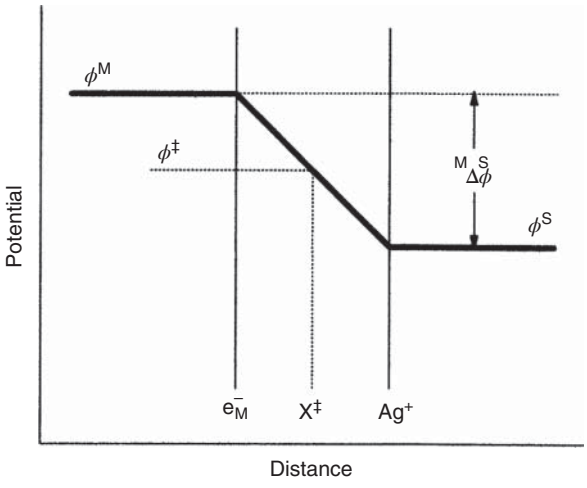


Figure 5.3 Charge transfer across the double layer. The reactant is at the potential ϕ^S, the product at ϕ^M, and the activated complex is at an intermediate position, $\phi^‡$, which is partway between ϕ^S and ϕ^M.

Now, what is the physics behind these equations? We note that the potential difference $\Delta\phi$ across the interphase affects the Gibbs free energy of the reaction. A fraction of this potential affects the Gibbs free energy of formation of the activated complex (i.e. the Gibbs free energy of activation of the reaction). This fraction is defined as β. When we conduct an experiment, we can control $\delta\Delta\phi$, but we have no control over $\delta\Delta\phi^‡$, which amounts to a fraction β of the former.

The important consequence of Eq. (4.7) for our present discussion is that the standard electrochemical Gibbs free energy of activation is a linear function of potential. The negative sign in Eq. (4.7) is applicable to anodic reactions, in which the rate is enhanced by increasing the potential in the positive direction, and the positive sign is applicable to cathodic reactions. In either case, the electrochemical Gibbs free energy of activation decreases when the potential is changed "in the right direction," namely positive for an anodic process and negative for a cathodic process.

It may also be appropriate to refer to the parameter β as an efficiency factor. Thus, we try to accelerate the rate of an electrochemical reaction by applying an amount of electrical energy$F\delta\Delta\phi$, but only a fraction β of this electrical energy is used to reduce the Gibbs free energy of activation, the rest being "lost" as far as our endeavor to accelerate the reaction rate is concerned.

Why is it then that we call it the symmetry factor? It could be argued that a value of $\beta = 0.50$ corresponds to a symmetrical situation, in which the activated complex is formed exactly halfway between the reactant and the product, as seen in Figure 5.3.

5.3.2 The Numerical Value of β

The determination of the numerical value of the symmetry factor β is a thorny problem in electrode kinetics. We might start with the conclusion; namely, that it is common practice to use the value of $\beta = 0.5$ (or, better, $\beta \approx 0.50 \pm 0.05$) in

the study of electrode reactions. It is hard to come up with a satisfactory theory showing why this should be so, but there seems to be good evidence that it is, at least in some experimental systems.

The most reliable data are from studies of hydrogen evolution on mercury cathodes in acid solutions. This reaction has been studied most extensively over the years. The use of a renewable surface (a dropping mercury electrode), our ability to purify the electrode material by distillation, the wide region of potential over which the Tafel equation is applicable, and the relatively simple mechanism of the reaction in this system – all combine to give high credence to the conclusion that the above value of β is reasonable, particularly for outer-sphere charge transfer reactions on mercury electrodes, in spite of the lack of solid theoretical evidence to support it.

5.4 The Marcus Theory of Charge Transfer

5.4.1 Outer-Sphere Electron Transfer

Electrode reactions can be divided into two major groups: those in which only charge is transferred across the interface, and those in which both charge and atomic mass are transferred. *Outer-sphere charge transfer* is a good example of the former, while metal deposition and dissolution are examples of the latter.

A typical outer-sphere charge transfer reaction is the ferri-ferro-cyanide redox couple

$$\left[\mathrm{Fe(CN)_6}\right]^{3-}_{\mathrm{soln}} + \mathrm{e}^-_{\mathrm{M}} \rightarrow \left[\mathrm{Fe(CN)_6}\right]^{4-}_{\mathrm{soln}} \tag{5.36}$$

This reaction occurs on the solution side of the interphase, and the electrodes serve only as source and sink for electrons. The important thing to note is that the close environment of the central ion is not changed significantly as a result of the transfer of charge. Admittedly, the configuration of the solvent molecules just outside the complex ion will change a little, but this is an effect on the "outer sphere." The inner sphere, that is, the six ligands around the central ion, remains intact. No chemical bond is broken, and none is formed, as a result of charge transfer.

Although Eq. (5.36) does not show any change in the inner sphere of the ion, it should be noted that some changes must occur. Thus, as a first approximation, the ligands can be viewed as dipoles interacting electrostatically with the central ion. If the charge on this ion changes as a result of electron transfer, surely the positions of the ligands must change. A detailed approach must also take into account the covalent bonding of the ligands to the central ion, considering that the wave function of the central ion changes upon addition or removal of a unit charge. Thus, transfer of an electron must change the total solvation energy of the ion: that associated with the inner sphere, as well as that associated with the outer sphere. Recognition of the importance of this energy change is a central issue in the Marcus theory of charge transfer, which is discussed further.

5.4.2 The Born–Oppenheimer Approximation

The essence of the *Born–Oppenheimer approximation* is that the electronic and nuclear motions (following charge transfer) are separable in time. This can be understood tentatively, if one considers that the $[Fe(CN)_6]^{3-}_{soln}$ ion is about 4×10^5 times heavier than an electron. It follows that the charge-transfer step presented by Eq. (5.36) should, in fact, be written in two steps, namely

$$\left[Fe(CN)_6\right]^{3-}_{soln} + e^-_M \rightarrow \left[Fe(CN)_6\right]^{\ddagger 4-}_{soln} \tag{5.37}$$

followed by

$$\left[Fe(CN)_6\right]^{\ddagger 4-}_{soln} \rightarrow \left[Fe(CN)_6\right]^{4-}_{soln} \tag{5.38}$$

In the first step, which is believed to occur very quickly, within a few femtoseconds (i.e., 10^{-15} seconds), an electron has been transferred, but the nuclei of the ligand, as well as the water molecules in the outer solvation shell, did not move yet. Although the positive charge on the central atom was reduced from +3 to +2, the configuration of the six ligands around the Fe^{2+} ion, as well as the hydration shell around it, are still in the equilibrium state corresponding to the Fe^{3+} ion. The superscripts on the complex $[Fe(CN)_6]^{\ddagger 4-}_{soln}$ in the above two equations indicate that the product of electron transfer is not in its most stable state. In the second step, which may typically take a few picoseconds (10^{-12} seconds), i.e. 2 or 3 orders of magnitude longer than the first step, the ligands around the newly formed Fe^{3+} ion are rearranged to their equilibrium position, corresponding to the lower charge. The variation of the Gibbs energy of the reactant and product along the reaction coordinate is shown schematically in Figure 5.4.

The total energy of the system is increased (because the species formed in Eq. (5.37) is not in its most stable state). The energy released when the product

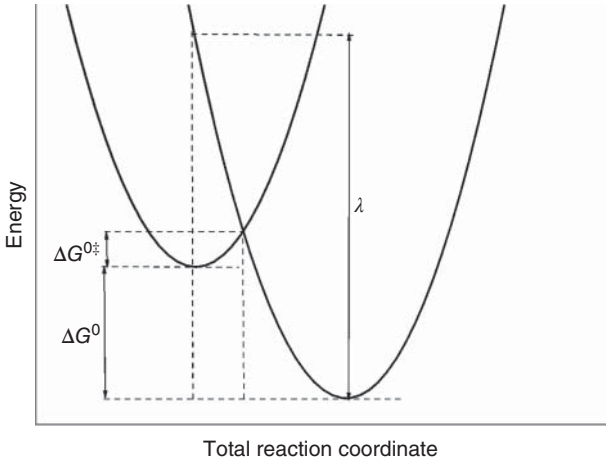


Figure 5.4 Schematic representation of "vertical" electron transfer, without movements of the nuclei, according to the Born–Oppenheimer approximation.

of electron transfer decays to its equilibrium state is referred to as *the solvent reorganization energy*, λ, which is of central importance in the Marcus theory of charge transfer.

This course of events presents no problem from the thermodynamic point of view: when a negative overpotential is applied to the system, which was initially at equilibrium, the change in the electrochemical Gibbs free energy is negative, and hence a driving force for the reaction to occur in the desired direction has been created. But, the widely different timescales for the two steps associated with Eqs. (5.37) and (5.38) do present a problem. The first step, which occurs in a few femtoseconds, creates an unstable intermediate, and hence increases the total Gibbs free energy of the system. However, there is no mechanism that could supply the extra energy in this very short time. It could take as long as 100–1000 fs for the unstable intermediate to relax to its stable final state. Consequently, the sequence of steps represented by Eqs. (5.37) and (5.38) would violate the law of conservation of energy! Thus, another route for charge transfer must be found. This problem is solved, in the framework of the Marcus theory, by assuming that charge transfer occurs only when the system is at the crossing point of the two curves shown in Figure 5.4. Since this point corresponds to the same energy level of the reactant and the product, transfer of an electron does not involve any change of the total Gibbs free energy of the system. Furthermore, it is assumed that the energy of the system is elevated to the crossing point of the two curves as a result of thermal fluctuations in the water, or any other solvent used.

5.4.3 The Calculated Energy of Activation

On the basis of the above physical model, we show the results of the *Marcus theory*, without showing the detailed calculations, which can be found in many textbooks of chemistry. Moreover, the equations shown here represent a first-order approximation to the full theory, intended only to explain the physical model involved.

The standard Gibbs free energy of activation for an outer-sphere charge-transfer process is given by

$$\Delta G^{0\ddagger} = \frac{\left(\lambda + \Delta G^0\right)^2}{4\lambda} \tag{5.39}$$

where λ is the solvent reorganization energy and ΔG^0 is the standard Gibbs free energy of the overall reaction, as seen in Figure 5.4. Equation (5.40) shows that solvent reorganization plays a major role in determining the rate of the outer-sphere electron-transfer processes. For a symmetrical electron transfer step taking place on the solution side of the interphase, where an electron is transferred from the reduced to the oxidized form of the same species, $\Delta G^0 = 0$, because the reactants and the products are the same, and Eq. (5.39) yields

$$\Delta G^{0\ddagger} = \frac{\lambda}{4} \tag{5.40}$$

It is not easy to calculate or experimentally determine the value of λ. Estimated values given in the literature indicate that it is higher in aqueous than in non-aqueous solutions of lower polarity.

5.4.4 The Value of β and its Potential Dependence

The results of the Marcus theory can be extended to outer-sphere electrode reactions by replacing the standard Gibbs free energies by the standard *electrochemical* Gibbs free energies, ΔG^0 and $\Delta\overline{G}^0$. Eq. (5.39) is thus rewritten as

$$\Delta\overline{G}^{0\ddagger} = \Delta G^{0\ddagger} - \beta F|\eta| = \frac{\left(\lambda + \Delta G^0 - F|\eta|\right)^2}{4\lambda} = \frac{\left(\lambda + \Delta G^0\right)^2}{4\lambda} - \frac{F|\eta|\left(2\lambda + 2\Delta G^0 - F|\eta|\right)}{4\lambda} \tag{5.41}$$

Combining Eqs. (5.40) and (5.41) one has

$$\beta F|\eta| = \frac{F|\eta|\left(2\lambda + 2\Delta G^0 - F|\eta|\right)}{4\lambda} \tag{5.42}$$

and the symmetry factor, β, is given by

$$\beta = \frac{1}{2} + \frac{\Delta G^0}{2\lambda} - \frac{F|\eta|}{4\lambda} \tag{5.43}$$

For charge transfer occurring across the electrode/solution interphase, ΔG^0 is not zero, since the reactant and the product are not identical. This could only lead to a value of $\beta \approx 0.50 \pm 0.05$ for high values of the solvent reorganization energy, or for a relatively short range of overpotentials, where the condition happens to apply.

$$\frac{\Delta G^0}{2\lambda} \approx \frac{F|\eta|}{4\lambda} \tag{5.44}$$

Moreover, according to Eq. (5.43), the Marcus theory predicts a potential-dependent symmetry factor, given by the expression

$$\frac{\partial \beta}{\partial |\eta|} = -\frac{F}{4\lambda} \tag{5.45}$$

which should lead to a deviation from linearity of the Tafel plot.

Certain experimental studies of outer-sphere charge-transfer processes seem to indicate that such a potential dependence of β does indeed exist, but this is based on measurements that were usually taken over a limited range of overpotential, so that the effect expected is very small, and the evidence is not compelling. On the other hand, highly accurate studies of the HER on mercury showed values of $\beta = 0.50 \pm 0.01$, strictly independent of potential over a range of more than six decades of current density. But, then, the HER is not an outer-sphere charge transfer process, so that this observation can neither prove nor contradict the Marcus theory.

5.5 Inner-Sphere Charge Transfer

5.5.1 Metal Deposition

In the previous section we discussed outer-sphere charge transfer, which is characterized by the following features:

(i) Both reactant and product are on the solution side of the interphase.
(ii) The charge is carried by electrons, which are the only species crossing the interphase.
(iii) There are no chemical bonds broken, and none are formed as a result of charge transfer. Indeed, the reactant and the product are very similar, except for their charge.

Inner-sphere charge transfer is very different from the outer-sphere model:

(i) The reactant is initially at the solution side of the interphase, but ends up on the surface of the substrate.
(ii) Chemical bonds (between the central metal ion and the hydration shell) are broken.
(iii) The hydrated ion crossing the interphase is typically 5–6 orders of magnitude heavier than the electron. As a result, the time required for it to cross the interphase is very long, on the atomic scale. It should therefore be obvious that the *inner* and *outer* charge transfer processes are physically different.
(iv) The energy of hydration of the metal ion is very high, on the scale of chemical bonds, and is expressed approximately by the equation

$$\Delta G_{\text{solv}} \approx 5 \times n^2 \,\text{eV} \tag{5.46}$$

where n is the charge of the hydrated cation. This corresponds to about 5 and 20 eV (482 and 1930 kJ) for $n = 1$ and 2, respectively. For comparison, the strongest chemical bond known is just below 10 eV. One may expect to find much lower values of the exchange current densities (hence, very slow kinetics) for inner-sphere than for outer-sphere charge transfer reactions. Moreover, j_0 for deposition of a divalent cation, such as Sn^{2+}_{soln}, should be even smaller than that of Ag^{+}_{soln}, considering Eq. (5.46). This is not borne out by experiment. Indeed, on comparing the highest rate constants for inner- and outer-sphere charge-transfer processes, comparable values are found. This presents an enigma, which has not been solved yet. How can the positive ion shed its hydration shell and allow deposition of the metal?

Deposition of a monovalent ion is usually written as

$$Ag^{+}_{\text{soln}} + e^{-}_{\text{M}} \rightarrow Ag^{0}_{\text{M}} \tag{5.47}$$

But, Eq. (5.47) represents the overall reaction, ignoring the steps along the way. These include the movement of the cation across the interphase, which is 5–6 orders of magnitude slower than the movement of the electron in the case of outer-sphere electron transfer. Studying Eq. (5.47) we note that the initial state is an ion in solution and an electron in the metal crystal. This could also be written as

$$Ag^{+}_{\text{soln}} + e^{-}_{\text{M}} \rightarrow Ag^{+}_{\text{M}} + e^{-}_{\text{M}} \tag{5.48}$$

Showing that the electron is in the metal, both in the initial and in the final state, it does not cross the interphase. Electroneutrality is maintained by movement of the electrons *in the external circuit*, from the anode to the cathode.

For the deposition of a divalent metal, a different route is assumed. The overall reaction is

$$Sn^{2+}_{soln} + 2e^{-}_{M} \rightarrow Sn^{0}_{M} \tag{5.49}$$

but since simultaneous two-electron transfer is usually not considered, and a monovalent Sn^{+}_{soln} is not known to exist in an aqueous solution, it is assumed that an intermediate is stabilized by adsorption on the surface

$$Sn^{2+}_{soln} + e^{-}_{M} \rightarrow Sn^{+}_{ads} \tag{5.50}$$

followed by

$$Sn^{+}_{ads} + e^{-}_{M} \rightarrow Sn^{0}_{M} \tag{5.51}$$

However, this assumption is also difficult to accept. In order to be stabilized by adsorption, the monovalent species has to be at the surface, in intimate contact with it. To do so, the unstable intermediate has to move across the interphase, similar to that discussed earlier for the deposition of Ag.

In conclusion, it is shown that the mechanism of metal deposition, and probably of other inner-sphere redox systems, must be different from that of outer-sphere processes, since electron transfer cannot be the correct path taking place. The species crossing the interphase is the relevant cation, not the electron. Unfortunately, a satisfactory mechanism has not yet been developed. However, further discussion of this issue is outside the scope of this textbook.

Further Reading

1 Bagotsky, V.S. (2006). *Fundamentals of Electrochemistry*, 2nd edn. New Jersey: Wiley.
2 Burstein, G.T. (2005). A hundred years of Tafel's equation: 1905–2005. *Corros. Sci.* 47: 2858–2870.
3 Gileadi, E. and Kirowa-Eisner, E. (2005). Some observations concerning the Tafel equation and its relevance to charge transfer in corrosion. *Corros. Sci.* 47: 3068–3085.
4 Newman, J. and Thomas-Alyea, K.E. (2004). *Electrochemical Systems*, 3rd edn. New Jersey: Wiley.
5 Bockris, J.O.'M., Reddy, A.K.N., and Gamboa-Aldeco, M. (2000). *Modern Electrochemistry*, Fundamentals of Electrodics, 2nd edn, vol. 2A. New York: Kluwer Academic Publishers.
6 Gileadi, E. (1993). *Electrode Kinetics – for Chemists, Chemical Engineers and Materials Scientists*. Wiley-VCH.
7 Tafel, J. (1905). Uber die polarisation bei kathodischer wasserstoffentwicklung. *Z. Phys. Chem.* 50: 641–712.
8 Tafel, J. and Naumann, K. (1905). Beziehungen zwischen kathodenpotential und elektrolytischer reduktionswirkung. *Z. Phys. Chem.* 50: 713–752.

6

Multistep Electrode Reactions

6.1 Mechanistic Criteria

6.1.1 The Transfer Coefficient, α, and its Relation to the Symmetry Factor, β

So far, we have discussed the current–potential relationship for a simple electrode reaction, which occurs in one step, in which a single electron is transferred. Usually it is assumed that there can be only one electron transferred in each elementary step, so that all single-step reactions involve the transfer of only one electron, although simultaneous two-electron transfer can occur in certain cases. Real systems are, however, more complex. Most electrode reactions occur in several steps, and the transition from reactants to products requires the transfer of several electrons, and in some cases also protons. When one considers a complex reaction, such as the oxidation of methanol in a fuel cell, the overall reaction is

$$CH_3OH + H_2O \rightarrow CO_2 + 6H^+ + 6e^-_M \tag{6.1}$$

which must occur in several electron- and proton-transfer steps, as it proceeds from reactants to products. We shall not discuss here the mechanism of this particular reaction. It is mentioned only to show the complexity of an electrode reaction, which happens to be of great interest in the context of the development of direct-methanol fuel cells. It is evident then, that the equations of electrode kinetics derived in Chapter 5 must be generalized to describe multistep electrode processes, which involve the transfer of several electrons, and often protons.

We recall that the Tafel relation was originally observed for the hydrogen evolution reaction (HER) on mercury and lead. We should therefore have an equation similar to Eq. (5.26) to describe the current–potential relationship. For a single-step anodic reaction we write

$$\frac{j}{j_0} = \exp\left(\frac{\beta F}{RT}\eta\right) \tag{5.26}$$

However, the symmetry factor β has been defined strictly for a single step and is related to the shape of the energy barrier and to the position of the activated complex along the reaction coordinate.

To describe a multistep process, β in Eq. (5.25) must be replaced by an experimental parameter, called the *transfer coefficient*, α. The value of α is determined

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

by the mechanism of the overall reaction, the position of the rate-determining step (RDS) in the reaction sequence. Unlike the symmetry factor, the sum of the transfer coefficients in the anodic and cathodic direction is not limited to unity. The equations showing the relationship between the current density in the anodic and cathodic direction are

$$\frac{j}{j_0} = \exp\left(\frac{\alpha_a F}{RT}\eta\right) \tag{6.2a}$$

and

$$\frac{j}{j_0} = \exp\left(-\frac{\alpha_c F}{RT}\eta\right) \tag{6.2b}$$

In the following text we shall just write α, for simplicity. The Tafel slope is then

$$b = \left(\frac{\partial|\eta|}{\partial \log j}\right) = \frac{2.3RT}{\alpha F} \tag{6.3}$$

These equations can be rearranged as

$$\alpha = \frac{2.3RT}{F}\left(\frac{\partial \log j}{\partial|\eta|}\right) \tag{6.4}$$

or

$$\alpha = \frac{2.3RT}{F}\frac{1}{|b|} \tag{6.5}$$

Either Eq. (6.3) or (6.5) can be considered to be a definition of the transfer coefficient. Equation (6.3) relates it directly to the measured quantity $(\partial\eta/\partial \log j)_{P,\,c,\,T}$, while Eq. (6.5) can be regarded as a formal definition, indicating that the transfer coefficient is simply the reciprocal Tafel slope in dimensionless form.

Although the definition of α and the difference between it and β should be quite clear from the preceding explanation, there has been a fair amount of confusion in the literature concerning these two parameters. It is therefore appropriate to state again the definitions of these quantities and to highlight their different physical meanings.

The symmetry factor, β, is a fundamental parameter in electrode kinetics. It must always be discussed with respect to a specific *single step* in a reaction sequence, and its value (which must be between zero and unity, by definition) is related to the shape of the energy barrier and to the position of the activated complex along the reaction coordinate. When the potential across the interface is changed (by application of an external current or potential), both the electrochemical Gibbs free energy of activation and the electrochemical Gibbs free energy of the reaction are altered. The symmetry factor represents the ratio between the changes in these quantities, as given by Eq. (4.9), namely

$$\beta \equiv \frac{(\delta\Delta\overline{G}^{0\ddagger}/\delta\Delta\phi)}{(\delta\Delta\overline{G}^{0}/\delta\Delta\phi)} = \frac{\delta\Delta\overline{G}^{0\ddagger}}{\delta\Delta\overline{G}^{0}} \tag{4.9}$$

In contrast, the transfer coefficient, α, is an experimental parameter obtained from the current–potential relationship. Just that and nothing more! It will be

shown later that the relationship between α and β depends on the mechanism of the reaction. The transfer coefficient is therefore *one of the parameters* that allow us to evaluate the mechanism of electrode reactions or to distinguish between different plausible mechanisms.

It cannot be overemphasized that one can measure only the transfer coefficient, α, not the symmetry factor, β. The latter can be inferred from the former by making a suitable set of assumptions. For example, for the HER on mercury, it is commonly *assumed* that $\alpha_c = \beta_c$. One often tends to refer in this case to the measurement of the symmetry factor, but even in this simple case, only the transfer coefficient can be measured, and the symmetry factor must be calculated from it (even if, in this case, "calculation" simply means assuming that they are equal).

There is an additional difference of great importance between β and α, which is often overlooked. The way in which β is defined requires that the sum of the symmetry factors in the anodic and cathodic directions must be unity; if it is β for the cathodic reaction, it must be $(1-\beta)$ for the anodic reaction, and vice versa (namely $\beta_c + \beta_a = 1$). The same is not true with respect to the transfer coefficient. To begin with, α is a parameter obtained from experiments. One must therefore *find* its value, not *assume* what it should be. When we wish to write a rate equation, such as Eq. (5.25), for a multistep reaction, the correct form will be Eqs. (6.2a) and (6.2b), in which

$$\alpha_c + \alpha_a = n/\nu \tag{6.6}$$

where n is the total number of electrons transferred and ν is the number of occurrences of the RDS in the electrode reaction, as written. Indeed, one can readily write mechanisms for which the transfer coefficient is greater than unity, as we shall see later.

6.1.2 Steady State and Quasi-Equilibrium

Consider a simple anodic reaction, such as the oxidation of Cl^- to Cl_2, which can occur in the following two steps:

$$Cl^-_{soln} \rightarrow Cl^0_{ads} + e^-_M \tag{6.7}$$

$$Cl^-_{soln} + Cl^0_{ads} \rightarrow Cl_2 + e^-_M \tag{6.8}$$

The concentration of the adsorbed intermediate can best be expressed in terms of the partial surface coverage, θ, which is defined as the surface concentration, Γ (mol cm^{-2}), divided by the maximum surface concentration, Γ_{max} (mol cm^{-2}):

$$\theta \equiv \frac{\Gamma}{\Gamma_{max}} \tag{6.9}$$

The net rate of formation of the adsorbed intermediate can be written as

$$\frac{\partial\theta}{\partial t} = k_1(1-\theta)c_{b,Cl^-} - k_{-1}\theta - k_2\theta\, c_{b,Cl^-} + k_{-2}(1-\theta)c_{b,Cl^-} \tag{6.10}$$

in which k_i are the potential-dependent electrochemical rate constants. This can be solved under steady-state conditions by setting $\partial\theta/\partial t = 0$. It is, however, easier to treat multistep reaction sequences in the framework of the *quasi-equilibrium*

assumption, as shown later. In considering Eq. (6.9) it should be understood that the term $(1-\theta)$ has the same role as c_b: it is the concentration of free sites on the surface, which act as reactant, just like c_b.

When a reaction occurs in several consecutive steps, the rate of all steps must be equal at steady state (otherwise, the system would not be at steady state). This rate is determined by the slowest step in the sequence, which we refer to as the RDS. In the preceding example, if k_1, the specific rate constant for the first step, shown by Eq. (6.7), is much smaller than that of the second step, k_2, the rate of the second step will effectively be limited by the supply of adsorbed intermediates by the rate of the first step. To visualize this situation, consider a potential applied to several resistors in series. The current is determined by the overall resistance, which is simply the sum of the resistances in series,

$$R = R_1 + R_2 + R_3 + \cdots + R_n \tag{6.11}$$

If R_1 is much larger than all others, it will be dominant, and one has approximately $R \approx R_1$.

To use this simple comparison, it must be realized that the rate constants of the various steps can be represented by the *inverse* of the corresponding resistances; the higher the resistance, the lower the rate constant, and vice versa. Thus, the overall effective rate constant is given by

$$\frac{1}{k} = \frac{1}{k_1} + \frac{1}{k_2} + \cdots + \frac{1}{k_n} \tag{6.12}$$

It is clear, therefore, that the overall rate constant is determined by the *lowest* individual rate constant. We note, in passing, that Eq. (6.12) is similar to Eq. (1.2), which correlates the overall current to the activation- and mass-transport-controlled currents:

$$\frac{1}{j} = \frac{1}{j_{ac}} + \frac{1}{j_L} \tag{1.2}$$

This should not be surprising, since the activation- and mass-transport processes always occur in series, and should combine to determine the overall rate in the same way as do several activation-controlled steps in series.

How do we know that there is a distinct RDS, one for which k_1 is much smaller than all other rate constants? The fact is that we do not know it a priori, and indeed one could envisage situations in which several rate constants in a sequence would be comparable in magnitude. However, it is recalled that in electrode kinetics, the rate constant is a function of potential, and in a reaction sequence, the rate constants of the different steps could depend differently on potential. Thus, in electrode kinetics, more than in other fields of chemical kinetics, one is likely to observe a single RDS, at least over a certain potential range. It is therefore common (and sensible) to treat the mechanisms of electrode reactions on the assumption that there is a well-defined RDS in each potential region studied, with possible transitions from one RDS to another, as the potential is changed.

We now proceed to the concept of quasi-equilibrium. If there is a distinct RDS in a reaction sequence, then *all other steps before and after it must be effectively at equilibrium*. This comes about because the overall rate is, by definition, very slow

compared to the rate at which each of the other steps could proceed by itself, and equilibrium in these steps is therefore barely disturbed. In order to better understand this, consider the specific example given earlier for chlorine evolution. Assume, for the sake of argument, that the values of the exchange current density, j_0, for steps (6.7) and (6.8) are 250 and 1 mA cm^{-2}, respectively. Now consider the application of a current density of 0.5 mA cm^{-2}. We can calculate the overpotential corresponding to each step in the sequence, from Eq. (5.7), namely

$$\frac{j}{j_0} = \frac{\eta F}{RT} \tag{5.6}$$

We find for step (6.7): $\eta = (RT/F)(0.5/250) = 0.05$ mV, and for step (6.8): $\eta = (RT/F)(0.5/1) = 12.8$ mV. The total overpotential is the sum of these values, namely $\eta = 12.85$ mV, of which only 0.05 mV (0.4%) is associated with the first step. Hence, we can consider this step to be effectively at equilibrium. In this way, we proceed to calculate the kinetic parameters for the reaction sequence, assuming that all steps other than the RDS are at equilibrium.

Why call it the *quasi*-equilibrium assumption? This has nothing to do with the mathematical treatment that follows; it is used only to soothe our conscience. Strictly speaking, the preceding assumption is self-contradictory. Equilibrium is defined as the state in which no net reaction takes place. There is an exchange reaction (the rate of which is represented in electrode kinetics by j_0), but the rates of the forward and backward reactions are equal. How can we say that a step in a reaction sequence is at equilibrium while it is proceeding at a finite rate in one direction? We use the term *quasi-equilibrium* in recognition of the fact that this is only an approximation that serves our purpose well, even as we remain fully aware of its logical limitations.

6.1.3 Calculation of the Tafel Slope

Equipped with the assumption of quasi-equilibrium, we can now proceed to calculate the Tafel slopes and some other kinetic parameters for a few very simple cases, to show how such calculations are made. In Chapter 7 we shall discuss the kinetics of several reactions that either have been important in the development of the theory of electrode kinetics or are of current practical importance.

Consider again the chlorine evolution reaction, and let us assume first that Eq. (6.7) is the RDS:

$$\mathrm{Cl^-_{soln}} \rightarrow \mathrm{Cl^0_{ads}} + \mathrm{e^-_M} \tag{6.7}$$

The rate of this step is given by

$$j = F k_1 c_{\mathrm{b,Cl^-}} (1 - \theta) \exp\left(\frac{\beta F}{RT} E\right) \tag{6.13}$$

This equation is applicable at high overpotentials, where the reverse reaction can be ignored (because this is the condition where a linear Tafel slope can be expected). Also, it is assumed in writing this equation (and all following kinetic equations) that mass-transport limitation is either negligible or has been corrected for in a quantitative manner. It should be noted that we use

here the symmetry factor, β, not the transfer coefficient, α. We do this because we are referring to a *specific step* in the reaction sequence and not the overall reaction.

Considering Eq. (6.13), we have good reason to assume that the partial surface coverage θ is very small, since the intermediate is formed in the RDS and is removed by the following step, which is much faster. Taking $\theta \ll 1$ and, hence, $(1-\theta) \approx 1$, we can rewrite Eq. (6.13) to a very good approximation as

$$j_1 = F k_1 c_{\mathrm{b,Cl^-}} \exp\left(\frac{\beta_a F}{RT} E\right) \tag{6.14}$$

For the overall reaction we can then write

$$2j_1 = nF k_1 c_{\mathrm{b,Cl^-}} \exp\left(\frac{\alpha_a F}{RT} E\right) = j_0 \exp\left(\frac{\alpha_a F}{RT} E\right) \tag{6.15}$$

It is important to understand the logic of the transition from Eq. (6.14) to (6.15). In the former we used the partial current density, j_1, for a step in the reaction sequence; in the latter we used the total current density, j. In the present case $j = 2j_1$, since for every electron transferred in the first step, another is transferred in the second step ($n = 2$). In addition, we write α_a instead of β_a, since Eq. (6.14) refers to an elementary step in the reaction sequence, while Eq. (6.15) refers to the overall reaction. In this particular case, we find that $\alpha_a = \beta_a$, but that is besides the point. It should also be noted here that the exchange current density is related to the rate constant, the concentration of the reactant, and the metal-solution potential difference at the reversible potential (since $E = E_{\mathrm{rev}} + \eta$):

$$j_0 = nF k c_{\mathrm{b,Cl^-}} \exp\left(\frac{\alpha F}{RT} E_{\mathrm{rev}}\right) \tag{6.16}$$

We can readily calculate the Tafel slope for this case, if we assign a numerical value to the symmetry factor β. This, as we said before, is commonly taken to be 0.50. The Tafel slope can then be obtained from

$$b = \left(\frac{\partial E}{\partial \log j_c}\right) = \frac{2.3RT}{\alpha F} = 118\ \mathrm{mV\ \ decade^{-1}\ \ (at\ 25^\circ C)} \tag{6.17}$$

We can now proceed to obtain the kinetic equation for the same reaction in a somewhat more complicated case, when the second step (Eq. (6.8)) is assumed to be the RDS, and the first step (Eq. (6.7)) is at quasi-equilibrium.

For quasi-equilibrium in step (6.7) we can write

$$k_1 c_{\mathrm{b,Cl^-}} (1-\theta) \exp\left(\frac{\beta F}{RT} E\right) = k_{-1}\theta \ \exp\left(-\frac{(1-\beta)F}{RT} E\right) \tag{6.18}$$

from which it follows that

$$\frac{\theta}{1-\theta} = K_1 c_{\mathrm{b,Cl^-}} \exp\left(\frac{F}{RT} E\right) \tag{6.19}$$

in which $K_1 = k_1/k_{-1}$ is the equilibrium constant. The absence of the symmetry factor β_a from this equation is not an error, and it does not depend on the numerical value of β_a. The symmetry factor is strictly related to the shape of the electrochemical Gibbs free energy barrier and the position of the activated complex along the reaction coordinate. Equation (6.19) describes an equilibrium that

is independent of the preceding considerations and is related only to the difference in the electrochemical Gibbs free energy between the initial and the final states.

Let us work out the kinetic equations for this case. Assuming that the second step, shown in Eq. (6.8), is the RDS

$$j_2 = F k_2 c_{\mathrm{b,Cl^-}} \theta \exp\left(\frac{\beta F}{RT} E\right) \tag{6.20}$$

in which we have to substitute the potential-dependent value of θ from Eq. (6.19). To simplify things, we consider the following two extreme cases:

1. For $\theta \ll 1$, we can write $(1-\theta) \approx 1$. Then, by combining Eqs. (6.19) and (6.20) we have

$$j = nFk_2 K_1 c_{\mathrm{b,Cl^-}} \exp\left(\frac{(1+\beta_\mathrm{a})F}{RT} E\right) \tag{6.21}$$

   from which it follows that, under these assumptions,

$$\alpha_\mathrm{a} = (1+\beta_\mathrm{a}) \quad b_\mathrm{a} = 2.3RT/(1+\beta_\mathrm{a})F \tag{6.22}$$

   and the Tafel slope turns out to be 39.3 mV per decade of current density at 25 °C.
2. When the surface coverage approaches unity and can no longer change significantly with potential, we can substitute $\theta = 1$ in Eq. (6.18), to obtain

$$j = nFk_2 c_{\mathrm{b,Cl^-}} \exp\left(\frac{\beta_\mathrm{a} F}{RT} E\right) \tag{6.23}$$

which leads to $\alpha_\mathrm{a} = \beta_\mathrm{a}$ and $b_\mathrm{a} = (2.3RT)/\alpha_\mathrm{a} F = 118$ mV. This result seems to be identical to that derived by assuming that the first step in the reaction sequence is rate-determining. There is one important difference, however. Equation (6.23) was derived for the case of high surface coverage, while Eq. (6.13) applies for low coverage, where $(1-\theta)$ can be replaced by unity. To distinguish between these possibilities, we must make independent measurements of the surface coverage.

The two limiting cases just discussed are approximately applicable for $\theta \leq 0.1$ and $\theta \geq 0.9$, respectively. In the intermediate region, one could readily solve the equation by substituting θ from Eq. (6.19) into Eq. (6.20); however, the result (besides being cumbersome) leads to a transfer coefficient that decreases gradually with increasing overpotential, from $\alpha_\mathrm{a} = 1 + \beta_\mathrm{a}$ to $\alpha_\mathrm{a} = \beta_\mathrm{a}$. This also means that the Tafel slope changes with potential. In other words, the Tafel *plot* (which is the plot of E versus $\log j$) is not linear in this intermediate range of coverage. It should be noted that the potential dependence of α_a shown here is due to the variation of the partial coverage with potential, *not* because β_a itself is potential dependent.

We have stated that Eq. (6.19) is equivalent to the Nernst equation. This can be proved in the following simple way. We write this equation in logarithmic form as

$$\log\left[\frac{\theta}{1-\theta}\right] = \log K_1 + \log c_{\mathrm{b,Cl^-}} + \frac{F}{2.3RT} E \tag{6.24}$$

which can be rearranged to

$$E = -\frac{2.3RT}{F}\log K_1 + \frac{2.3RT}{F}\log\left[\left(\frac{1}{c_{\mathrm{b,Cl^-}}}\right)\left(\frac{\theta}{1-\theta}\right)\right] \tag{6.25}$$

The fractional coverage, θ, is proportional to the concentration of the product in this reaction, Cl_{ads}, and $(1-\theta)$ is proportional to the concentration of one of the reactants (free sites on the surface), making the second term in Eq. (6.25) a typical ratio of concentrations of products to reactants. Moreover, we know from thermodynamics that the equilibrium constant is related to the standard Gibbs free energy of the reaction, and the latter is related to the standard potential:

$$\frac{2.3RT}{F}\log K = -\frac{\Delta G^0}{F} = E_\theta^0 \tag{6.26}$$

Substituting in Eq. (6.24) we thus have

$$E = E_\theta^0 + (2.3RT/F)\log\left[\left(\frac{1}{c_{\mathrm{b,Cl^-}}}\right)\left(\frac{\theta}{1-\theta}\right)\right] \tag{6.27}$$

where E_θ^0 is the standard potential for a reaction involving the formation of an adsorbed species by charge transfer, such as described by Eq. (6.7). We note that the standard state for the adsorbed species is chosen as $\theta = 0.5$. Hence, the term $\theta/(1-\theta)$ will be unity. Equation (6.27) is the Nernst equation, written for a single-electron transfer. There is nothing unusual in this result, of course. We treated Eq. (6.7) as an electrochemical equilibrium and obtained the correct equation for an electrochemical reaction at equilibrium. The point is that the Nernst equation is commonly derived from thermodynamic considerations. It is reassuring to be able to start from kinetic equations and reach the same result, in the limit where equilibrium can be assumed.

6.1.4 Reaction Orders in Electrode Kinetics

The *reaction order* is defined in chemical kinetics by the partial derivative

$$\rho = \left(\frac{\partial \log\, \nu}{\partial \log\, c_i}\right)_{c_{j\neq i}} \tag{6.28}$$

which measures the dependence of the reaction rate, ν, on the concentration of one species in solution, with the concentrations of all other species (as well as the temperature and pressure) maintained constant.

In electrode kinetics, the reaction order is defined in a similar manner, but, in addition to keeping the temperature and the pressure constant, a constant potential is maintained. As a result, there are two reaction orders in electrode kinetics, one taken at constant *potential*

$$\rho_1 = \left(\frac{\partial \log j}{\partial \log\, c_i}\right)_{E,c_{j\neq i}} \tag{6.29}$$

and the other at constant *overpotential*

$$\rho_2 = \left(\frac{\partial \log j}{\partial \log\, c_i}\right)_{\eta,c_{j\neq i}} \tag{6.30}$$

It is important to distinguish between these two parameters, since the reversible potential changes as we change the concentration of the reactant, and the overpotential can thus change while the applied potential remains constant.

In Eq. (6.29), the potential, E, is measured with respect to a fixed reference electrode. It does not matter which reference electrode is used, as long as it is the same throughout the experiment. Actually, we would like to obtain the reaction order while keeping the metal–solution potential difference, ${}^{S}\Delta^{M}\phi$, constant. This is impossible, since ${}^{S}\Delta^{M}\phi$ cannot be measured, as explained in detail in Chapters 1 and 2. Nevertheless, the difference between the *absolute* metal–solution potential and the *measured* potential is only a constant! Thus, by keeping the potential E constant, we can be assured that the metal–solution potential difference ${}^{S}\Delta^{M}\phi$ is also maintained constant, even though its value is not known.

We recall that the current density is a very sensitive measure and influences the rate of an electrochemical reaction. It is therefore quite easy to determine the current–potential relationship without causing a significant change in the bulk concentration of either reactants or products. Thus, measurements in electrode kinetics are conducted effectively under *quasi-zero-order* kinetic conditions. It would be wrong to infer from this that electrode reactions are independent of concentration. To determine the concentration dependence (i.e. the reaction order), one must obtain a series of j/E or j/η plots, as shown in Figure 6.1a, and derive from them the dependence of $\log j$ on $\log c_i$ at different potentials, as shown in Figure 6.1b. The slopes in Figure 6.1b yield the parameter $\rho_{1,}$ since it is measured at constant potential E. Here, and in all further equations, we shall assume that T, P, and the concentration of all species in solution, $c_{j\neq \mathrm{i}}$ other than the one being studied, are kept constant, in order to permit us to write the equations in a more concise form.

Which of the two reaction-order parameters should one prefer? When measurement is made with a constant reference electrode (e.g. silver/silver chloride), the reaction order at constant potential, ρ_1, is obtained. If, however, an indicator-type reference electrode is employed (e.g. a reversible hydrogen electrode, often used in the study of the HER), the overpotential can be kept constant, and the parameter ρ_2 is the one directly obtained. In either case, both parameters can be calculated from the same experimental data, since the dependence of the reversible potential on concentration is well known.

To show the difference between these two reaction-order parameters, we return to the equations derived in Section 6.1.3 for the chlorine-evolution reaction. Consider the case in which the first charge-transfer step (see Eq. (6.7)) is rate-determining. Equation (6.14) in logarithmic form is

$$\log j = \log(nFk_1) + \log c_{\mathrm{b,Cl^-}} + \left(\frac{\alpha_{\mathrm{a}} F}{RT} E\right) \tag{6.31}$$

hence,

$$\rho_1 = \left(\frac{\partial \log j}{\partial \log c_{\mathrm{b,Cl^-}}}\right)_E = 1 \tag{6.32}$$

To calculate ρ_2, the reaction order at constant overpotential, we substitute $E = E_{\mathrm{rev}} + \eta$ in Eq. (6.31) and express the reversible potential in terms of the

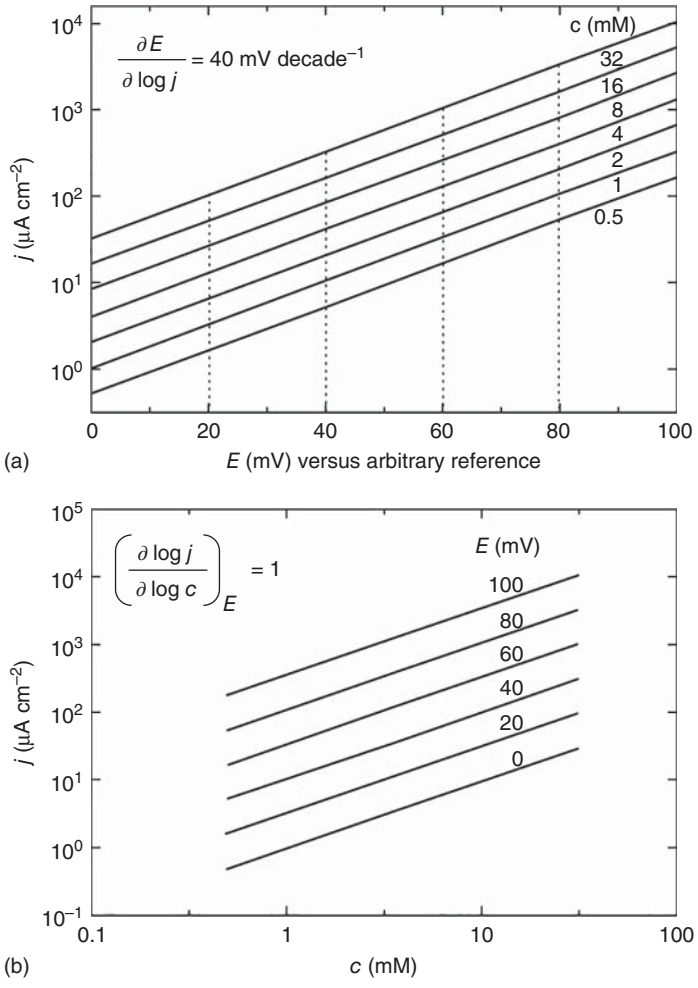


Figure 6.1 Calculation of the reaction order from Tafel plots. (a) Tafel lines for a series of concentrations of the electroactive species. (b) Reaction-order plots derived from (a) for different values of the potential.

Nernst equation, namely

$$E_{\text{rev}} = E^0 + \frac{2.3RT}{2F} \log \frac{c_{\text{Cl}_2}}{(c_{\text{b,Cl}^-})^2} \tag{6.33}$$

Assuming that the solution is saturated with respect to Cl_2, we can rewrite Eq. (6.31) as

$$\log j = \log(nFk_1) + \log c_{\text{b,Cl}^-} + \frac{\alpha_{\text{a}} F}{RT} E^0 - \alpha_{\text{a}} \log c_{\text{b,Cl}^-} + \frac{\alpha_{\text{a}} F}{RT} \eta \tag{6.34}$$

Differentiating with respect to $\log c_{\text{b,Cl}^-}$ while keeping η constant yields

$$\rho_2 \equiv \left(\frac{\partial \log j}{\partial \log c_{\text{b,Cl}^-}} \right) = 1 - \alpha_{\text{a}} \tag{6.35}$$

We note that the reaction order at constant overpotential is a little more complex and depends also on the transfer coefficient α_a. One may be tempted to obtain both ρ_1 and ρ_2 from the experimental data, in order to calculate the transfer coefficient from the difference. This can be done, in principle, but the measurement of reaction orders is inherently much less accurate than the measurement of the Tafel slope, hence, this approach is not expedient, unless the transfer coefficient cannot be measured directly.

In the literature, one often encounters two additional types of reaction orders, which are actually derived from the two we have just discussed. The first relates to the variation of the *exchange current density* with concentration, namely $(\partial \log j_0/\partial \log c)$. It is easy to see that this parameter is equal to ρ_2, since by using j_0 one has actually specified a constant overpotential of $\eta = 0$. This parameter is useful for the study of fast electrode reactions, where measurements can be taken only at low overpotentials (as a result of mass-transport limitation), and the Tafel region is not experimentally accessible. The second reaction-order parameter occasionally used is $(\partial E/\partial \log c_i)_{c_{j\neq i}}$. However, applying the rule of partial derivatives, we can write

$$\left(\frac{\partial E}{\partial \log c}\right)_j \left(\frac{\partial \log c}{\partial \log j}\right)_E \left(\frac{\partial \log j}{\partial E}\right)_c = -1 \tag{6.36}$$

Hence,

$$\left(\frac{\partial E}{\partial \log c}\right)_j = -\rho_1 b \tag{6.37}$$

There is nothing wrong with using any of these reaction-order parameters, as long as we remember that only two of them (any two, of course) are independent, and the third can always be derived from the other two.

6.1.5 The Effect of pH on Reaction Rates

The influence of pH on reaction rates may be looked upon as just another concentration effect, which can be dealt with in terms of the reaction orders just discussed. It merits special attention, however, for two reasons: (i) because it allows us to change the concentration of the reactant (H_3O^+ or OH^-) over many orders of magnitude, without encountering solubility limitations at the high concentration end, and mass transport limitations at the low concentration end (as long as the solution is buffered), and (ii) because in aqueous solutions, the solvent itself can be the reactant or the product in the reaction being studied.

However, a few words of caution might be appropriate in regard to the use of this concept, particularly for readers who did not major in chemistry. The pH of an aqueous solution is formally defined as

$$\mathrm{pH} \equiv -\log a_{H_3O^+} \tag{6.38}$$

in which $a_{H_3O^+}$ is the activity of the hydronium ion. Single-ion activity cannot be measured, but in dilute solutions, the activity can be approximated by the concentration. This approximation is so often made that one may forget that it applies only to dilute solutions, where the physical properties of the solvent are

not perceptibly affected by the addition of solute. In contrast, in a 7.0 M solution of KOH the pH is about 16, far from the value of 14.85 calculated from Eq. (6.38), employing concentration instead of activity. Also, if one uses a mixed solvent (e.g. methanol and water), the concept of pH becomes somewhat more complicated, and must be employed with caution. In such cases, even if the pH can be measured by a conventional method (e.g. with a glass electrode), the result may represent quite different concentrations of H_3O^+ ions in different solvent mixtures.

When the temperature is changed, the equilibrium constant for the dissociation of water is also changed. As a result, pH 7.0 represents the point of neutrality only at 25 °C. At higher temperatures, the point of neutrality moves to lower pH values. Similarly, in mixed solvents the point of neutrality is not necessarily at pH 7.0.

Let us now return to the effect of pH on electrode kinetics, using concentrations instead of activities. Consider the HER, and assume that it proceeds in the following two steps, with the second step being rate-determining:

$$H_3O^+ + e_M^- \rightleftharpoons H_{ads} + H_2O \tag{6.39}$$

$$2H_{ads} \xrightarrow{\text{RDS}} H_2 \tag{6.40}$$

For equilibrium in the first step we can write

$$\frac{\theta}{1-\theta} = K_1 c_{b,H_3O^+} \exp\left(-\frac{F}{RT}E\right) \tag{6.41}$$

This is identical to Eq. (6.19), derived for the first step in chlorine evolution, except for the change in the sign of the exponent, which is necessary because we are now dealing with a cathodic reaction.

For the RDS we write

$$j_2 = F k_2 \theta^2 \tag{6.42}$$

This is an interesting case, which may need some clarification. Equation (6.40), as written, is not an electrochemical step in the sense that it does not involve charge transfer. Nevertheless, we are justified in expressing the rate of this step in terms of current density, since every time step (6.40) occurs, two electrons must have been transferred in the preceding step.

If we assume low coverage ($\theta \ll 1$), we can substitute θ from Eq. (6.41) into Eq. (6.40), to obtain

$$j_2 = nF K_1^2 k_2 (c_{b,H_3O^+})^2 \exp\left(-\frac{2F}{RT}E\right) \tag{6.43}$$

from which we obtain directly

$$\rho_1 = \left(\frac{\partial \log j}{\partial \log c_{b,H_3O^+}}\right)_E = -\left(\frac{\partial \log j}{\partial \text{pH}}\right)_E = 2 \tag{6.44}$$

Also, since we can write the Nernst equation in this case as

$$E_{rev} = E^0 + \left[\frac{2.3RT}{2F}\right] \log \left[\frac{(c_{b,H_3O^+})^2}{P_{H_2}}\right] \tag{6.45}$$

we have, from Eq. (6.43), taking $P_{H_2} = 1$:

$$\log j_2 = \log(nFK_1^2 k_2) + 2\ \log(c_{b,H_3O^+}) - 2\left(\frac{2F}{2.3RT}E^0\right) - 2\ \log(c_{b,H_3O^+}) - 2\left(\frac{2F}{2.3RT}\right)\eta \tag{6.46}$$

Hence,

$$\log j_2 = \log(nFK_1^2 k_2) - \frac{2F}{2.3RT}(E^0 + \eta) \tag{6.47}$$

We note that, in this particular case, the reaction order at constant overpotential, ρ_2, becomes independent of pH. When the pH is lowered, the effect of increasing concentration of the hydronium ion is exactly compensated for by the influence of the variation of the reversible potential on the reaction rate.

It is interesting to note that the symmetry factor β does not appear in any of these equations. This is because the RDS assumed here does not involve charge transfer. The current depends indirectly on potential, through the potential dependence of the fractional coverage, θ. The transfer coefficient is $\alpha_c = 2.0$ exactly, as can be seen in Eq. (6.44), and this corresponds to a Tafel slope of $b_c = 29.5$ mV at room temperature.

Note that the transfer coefficient obtained here is not, in any way, related to the symmetry factor. It arises from the quasi-equilibrium assumption and should therefore be a true constant, independent of potential, as long as the assumptions leading to Eq. (6.45) are valid.

6.1.6 The Enthalpy of Activation

To discuss the enthalpy of activation in electrode kinetics, we make use of the fundamental rate equation

$$j = nFc_i \frac{kT}{h} \exp\left(-\frac{\Delta G^{0\ddagger}}{RT}\right) \exp\left(\frac{\alpha_a F}{RT}E\right) \tag{6.48}$$

which is similar to Eq. (5.21), except that it is written for a multistep anodic reaction. We split the Gibbs free energy of activation into its enthalpic and entropic parts, and rewrite Eq. (6.48) in logarithmic form as

$$\log j = \log\left(nFc_i \frac{kT}{h}\right) - \frac{\Delta H^{0\ddagger}}{2.3RT} + \frac{\Delta S^{0\ddagger}}{2.3R} + \frac{\alpha_a F}{2.3RT}E \tag{6.49}$$

from which it follows that a plot of $\log j$ versus $1/T$ should be a straight line with a slope of

$$\left[\frac{\partial \log j}{\partial(1/T)}\right] = -\frac{\Delta H^{0\ddagger} - \alpha_a FE}{2.3R} = -\frac{\Delta \overline{H}^{0\ddagger}}{2.3R} \tag{6.50}$$

This is similar to the usual treatment in chemical kinetics, except that the enthalpy of activation is found to be a function of potential, as shown in Figure 6.2.

Often, reference is made in the literature to the *energy* of activation, instead of the *enthalpy* of activation. It follows from elementary thermodynamics that

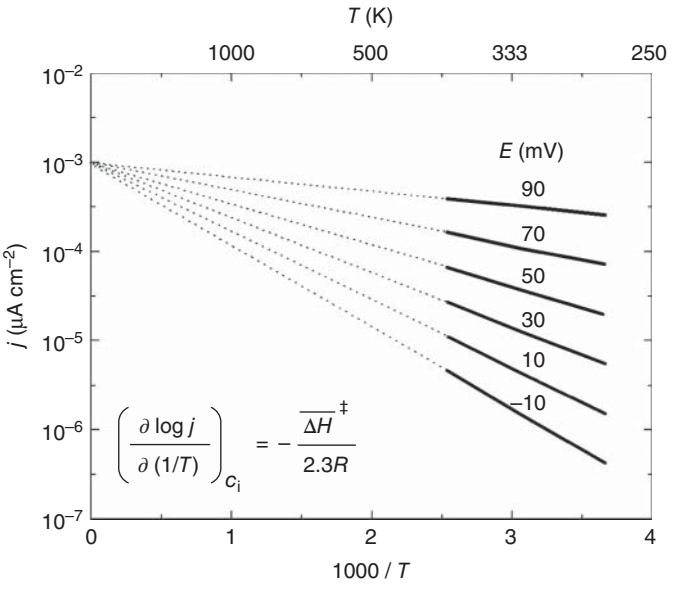


Figure 6.2 Schematic Arrhenius plots, showing the variation of the apparent electrochemical enthalpy of activation as a function of potential. All lines extrapolate to the same point at $1000/T = 0$, i.e. at infinite temperature.

the former applies if the reaction is conducted at constant volume, whereas the latter is applicable to conditions of constant pressure. In aqueous solutions, the difference between the two is negligible, and when measurements are made at ambient pressure, the terms can be used interchangeably.

To be exact, it should be noted that the frequency term kT/h in Eq. (6.48) contains a temperature term. This should lead to a nonlinear dependence of $\log j$ versus $1/T$. The effect is small, however, and can be neglected in the narrow temperature range accessible to experiments in aqueous solutions.

An inherent source of uncertainty in the calculation of the enthalpy of activation, unique to electrochemistry, is related to the temperature dependence of the potential of the reference electrode. Thus, in order to obtain $\Delta\overline{H}^{0\ddagger}$, we determine $\log j$ versus $1/T$ at a constant metal–solution potential difference, $\Delta\phi$. Now, at any given temperature, $\Delta\phi$ is constant, as long as the potential with respect to a given reference electrode is constant. When the temperature is changed, this is no longer true, since the metal–solution potential difference at the reference electrode has changed by an unknown amount.

There are two ways to approach this problem: measurement can be conducted *isothermally* (i.e. keeping the working and the reference electrodes always at the same temperature) or *non-isothermally* (i.e. keeping the temperature of the reference electrode constant, while that of the working electrode is scanned). In the latter case, the reference and working electrode compartments are connected through a salt bridge, and there must be a temperature gradient somewhere along this bridge, causing a small thermal-junction potential.

There is some merit to each of these methods, and both have been used. In isothermal mode, we estimate the change in the value of $\Delta\phi$ with T for the reference electrode. In non-isothermal mode, we estimate the additional potential drop generated in the salt bridge as a result of the thermal gradients. Neither can be measured directly. Consequently, there is always some uncertainty in the value of the enthalpy of activation of electrode reactions.

We note (see Figure 6.2) that the enthalpy of activation decreases with potential (or rather with increasing overpotential), and its rate of change with potential is proportional to the transfer coefficient, α_a

$$\Delta\overline{H}^{0\ddagger} = \Delta H^{0\ddagger} - \alpha_a FE \tag{6.51}$$

Thus, α_a could, in principle, be obtained from Eq. (6.51), but since direct measurement of α_a from the current–potential relationship is much more accurate, it is better to introduce its value, determined from the Tafel slope, into Eq. (6.51) in order to test the accuracy of the measured values of the electrochemical enthalpy of activation, $\Delta\overline{H}^{0\ddagger}$.

Further Reading

1 Guidelli, R., Compton, R.G., Feliu, J.M. et al. (2014). Defining the transfer coefficient in electrochemistry: an assessment (IUPAC technical report). *Pure Appl. Chem.* 86: 245–258.

2 Andriiko, A.A., Andriyko, Y.O., and Nauer, G.E. (2013). Many-electron electrochemical processes: reactions in molten salts. In: *Room-Temperature Ionic Liquids and Ionic Solutions*. Berlin, Heidelberg: Springer-Verlag.

3 Gileadi, E. and Kirowa-Eisner, E. (2005). Some observations concerning the Tafel equation and its relevance to charge transfer in corrosion. *Corros. Sci.* 47: 3068–3085.

4 Newman, J. and Thomas-Alyea, K.E. (2004). *Electrochemical Systems*, 3rd edn. New Jersey: Wiley.

5 Bockris, J.O.'M., Reddy, A.K.N., and Gamboa-Aldeco, M. (2000). *Modern Electrochemistry*, Fundamentals of Electrodics, 2nd edn, vol. 2A. New York: Kluwer Academic Publishers.

6 Gileadi, E. (1993). *Electrode Kinetics – for Chemists, Chemical Engineers and Materials Scientists*. Wiley-VCH.

7

Specific Examples of Multistep Electrode Reactions

7.1 Experimental Considerations

7.1.1 Multiple Processes in Parallel

In all the discussions so far, it has been tacitly assumed that we know the reaction taking place and that *only one reaction* is occurring in the potential range of interest. Unfortunately, this is not always the case. In the electroplating industry, for example, one must specify the so-called *faradaic efficiency* (also called the *current efficiency*), which is the fraction of the current utilized for metal deposition (the rest is usually taken up by hydrogen evolution). The measured current in itself does not yield any information regarding the fraction of the current supporting each reaction, of course. If two or more reactions occur simultaneously, the current–potential relationship can become rather complicated, since the different partial currents may depend on the potential in different ways. In the case of metal deposition, one can obtain the partial current density by measuring the mass of the metal deposited or the volume of hydrogen evolved (or both, for double checking). However, in that case, the high sensitivity and ease of determination of the rate of electrode reactions, offered by the simple measurement of the current density, is lost.

In other cases, it may not be obvious how far the reaction can proceed under a given set of conditions. Oxygen, for instance, can be reduced either to hydrogen peroxide ($2e^-$ transfer) or to water ($4e^-$), depending on the nature of the electrode, the composition and purity of the solution, and the range of the potential studied. There could also be a region in which both reactions occur simultaneously, and their relative rates depend on the potential. In such cases, analysis of the *j/E* relationship could become rather complicated, and extracting the mechanistic parameters may not be possible.

Another tacit assumption is that the surface of the electrode remains unchanged during the experiment. Changes in the catalytic activity of the surface (which would change the specific rate constant of the reaction) could occur during metal deposition or dissolution, particularly when alloy deposition is concerned, where the atomic composition of the surface could change during the course of measurement, or the surface roughness could increase or decrease. Adsorption of impurities could cause major changes in the catalytic activity of

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

the electrode, particularly if it is nonuniform and contains active sites where most of the reaction takes place. This problem can be partially alleviated when a renewable surface, such as a dropping mercury electrode, is used as shown later. On solid electrodes it is practically impossible to maintain the desired high level of surface purity, unless the surface is held in a state where adsorption of impurities is inherently prevented. For example, oxygen evolution on platinum can be conducted without interference from most impurities, since the oxide-covered surface resists contamination by most impurities. On the other hand, low-frequency measurements of electrochemical impedance spectroscopy (EIS), which will be discussed in Chapter 15, take a long time, and the results obtained may be corrupted by variations of the surface during the course of the measurement.

7.1.2 The Level of Impurity that can be Tolerated

The above arguments give rise to the need for extreme conditions of purity in electrochemical measurements. Why is high purity necessary and how clean should the system be in order that reliable results will be obtained? While it may be a good idea to conduct *any* experiment in a clean system, it must be remembered that purity has a price, and for extreme purity the price may be quite high, in terms of both the cost of electrode materials and other chemicals, and the time and effort required for each experiment. Impurities can be divided into two groups: those that are electroactive in the potential range of interest, and those that may interfere with measurement by adsorbing on the surface and poisoning it (a more general and perhaps less ominous, phrase would be "changing its catalytic activity").

For the former group, the allowed level of impurity is relatively easy to assess. It should be several orders of magnitude less than the concentration of the material being studied, so that the mass-transport-limited current density at which the impurity can react will be smaller than the smallest current we wish to measure.

Let us clarify this argument by an example: the need to remove oxygen from solution during the study of the HER on mercury. The concentration of oxygen in a dilute aqueous solution at equilibrium with air is about 0.25 mM (8 ppm). For a dropping mercury electrode, which we renew once a second, this should yield a limiting current of about 3 μA. This is the current calculated for the four-electron reduction of oxygen to water, which is the reaction taking place at high negative overpotentials with respect to oxygen reduction, over the range in which hydrogen evolution is studied (see the Ilkovic Eq. (4.17)). For the HER on mercury, one typically tries to obtain the *j/E* relationship for current densities from 1×10^{-2} to 1×10^{4} μA cm^{-2}. It is necessary, therefore, to reduce the concentration of oxygen by 2–3 orders of magnitude, so that the limiting current for this reaction will be lower than the lowest current we wish to measure. This can be done by bubbling purified nitrogen or argon though the solution for some minutes before the experiment begins. The partial current for oxygen reduction and hydrogen evolution are shown in Figure 7.1, for two concentrations of oxygen. This figure shows that reducing the concentration of oxygen by a factor of 250 should extend the range over which the HER can be studied by about 0.28 V.

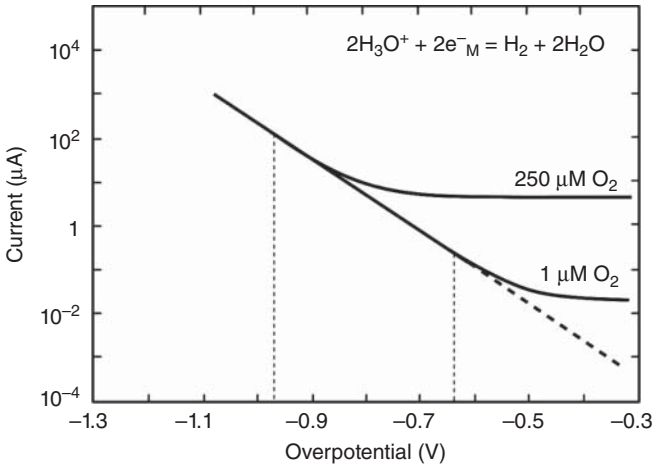


Figure 7.1 Current calculated for hydrogen evolution on a DME, assuming $I_0 = 3.3 \times 10^{-12}$ A ($j_0 = 1 \times 10^{-10}$ A cm^{-2}) in the presence of two concentrations of dissolved oxygen. Dotted line shows the current that would be observed in the total absence of oxygen.

It should be added that the experimental situation could be better than that indicated in this figure, because it is the *fluctuation* in the background current, not its absolute value, which determines the lowest current for hydrogen evolution that can be reliably evaluated. Thus, if the experiment is conducted on an RDE, where the limiting current for oxygen reduction is stable and can be measured accurately, the range over which the HER can be studied could be extended, by subtracting the oxygen–reduction current from the total current measured.

Now, consider the effect of an impurity that is adsorbed on the surface and alters its catalytic properties. We recall that it takes a very small amount of material (about 2×10^{-9} mol cm^{-2}) to form a monolayer. Moreover, very often a fraction of a monolayer is enough to change the properties of the surface significantly. In a typical experiment, there will be about 25 cm^3 of solution per square centimeter of electrode area. If the concentration of impurity is 1.0 μM, the total amount of impurity in solution will be 25 nmol, for each square centimeter of surface area, enough to form 10–20 molecular layers, depending on the size of the species being adsorbed. This is too high, even if we accept the fact that only a fraction of these impurity molecules reach the surface and adhere to it during measurement. Thus, a higher level of purity is required, if the solution is to be in contact with the electrode for a long time. A relatively simple way to ease the requirement for very high purity is to decrease the volume of the solution per unit surface area. Reducing this ratio to 1 cm^3 of solution per 1 cm^2 of surface area can readily be achieved. Moreover, thin-layer cells have been built with a gap of 10 μm, corresponding to a volume of 1×10^{-3} cm^3 per square centimeter. With an impurity level of 1.0 μM, this corresponds to a mere 1×10^{-12} mol cm^{-2} of surface area, namely 0.05% of a monolayer. Thus, by employing a thin-layer cell, one could relax the purity requirement by 1 or 2 orders of magnitude and still maintain a sufficiently clean surface during measurement.

The other aspect of allowed impurity levels relates to the maximum rate of adsorption, as compared to the duration of the experiment, or rather to the time interval between successive renewals of the surface. We shall explain this in relation to studies on the hanging dropping mercury electrode, which is typically renewed every second, although the argument can be applied also to solid electrodes, under certain favorable conditions, as we shall see.

The rate of adsorption of an impurity may be kinetically controlled, in which case it depends on the specific system being studied. This cannot be discussed in general terms, but we can calculate *the maximum rate of adsorption*, which is controlled by mass transport, as a function of the concentration of impurity. This rate is given by the flux of the impurity molecules reaching the surface. The logic behind this argument is very simple. If every molecule of impurity reaching the surface is instantaneously adsorbed, the concentration of these molecules in the solution nearest to the surface (i.e. at $x = 0$) will be zero. The rate of adsorption will then depend on the rate of supply of molecules to the surface, that is, it will be totally mass-transport-controlled. This situation is similar to the case in which an electroactive material is oxidized or reduced at the limiting current. In the case of adsorption, charge transfer may not take place, but the adsorbed molecules are removed from the solution, just as if they had reacted at the surface. Clearly, this is a worst-case scenario, since the rate of diffusion is the maximum rate for a given concentration, and the assumption that each species reaching the surface is adsorbed may not be valid. Thus, the degree of coverage could be less than that calculated in this manner, but it could not exceed it.

In the case of the dropping mercury electrode, the limiting current for a one-electron reduction is typically about 3 $\mu A\,mM^{-1}$, equal to a current density of about 100 $\mu A\,cm^{-2}\ mM^{-1}$ for a typical surface area of 0.03 cm^2. The flux of the impurity reaching the surface, in units of $mol\,cm^{-2}\ s^{-1}$, is equal to the diffusion-limited current density, divided by the charge per mole, nF. Assuming, as before, an impurity concentration of 1.0 μM, we obtain a flux of

$$\frac{j_d}{F} = \frac{(1 \times 10^{-4} A\,cm^{-2}\,mM^{-1})(1 \times 10^{-3}\,mM)}{9.65 \times 10^4 C\,mol^{-1}} \approx 1 \times 10^{-12} mol\,cm^{-2}\,s^{-1} \tag{7.1}$$

Considering that a monolayer amounts to about 2 $nmol\,cm^{-2}$, it would take about two thousand seconds to deposit a monolayer of impurity. Since the drop is renewed every second, the maximum coverage by an impurity that exists in solution at a concentration of 1.0 μM cannot exceed $\theta = 5 \times 10^{-4}$.

The foregoing calculation shows that one can relax the requirements for purification substantially. Thus, allowing a tenfold increase in the impurity level would still limit the maximum impurity coverage during the lifetime of a drop to $\theta \leq 5 \times 10^{-3}$.

These conditions can be realized in the simplest way on a dropping mercury electrode, but they are not totally limited to liquid metals. The surface of a solid electrode could also be cleaned or effectively renewed periodically, for example, by a suitable series of pulses of potential, by mechanical abrasion *in situ*, or by sudden heating with a laser pulse.

On the other hand, when a solid electrode is employed and the surface is not renewed periodically, the electrode may be in contact with the solution for 10^3–10^4 seconds, allowing ample time for the impurity to diffuse to the surface and be adsorbed on it. The requirements for solution purity are much more stringent in this case, of course.

So far, we have discussed only the *maximum* extent and rate of adsorption. This represents the worst case, and one cannot go wrong if the preceding requirements are satisfied. In general, however, adsorption depends on the potential. This can best be treated in the context of electrosorption of organic materials, discussed in Chapter 12. We shall limit the discussion here by saying that adsorption depends mostly on the charge density on the metal. Impurities that carry a negative charge will be adsorbed when the excess charge density on the electrode surface is positive, and vice versa. It is also true, although not obvious from first principles, that neutral species tend to be adsorbed mainly in the region in which the excess charge density is the lowest. The so-called *potential of zero charge,* E_{pzc}, where the excess charge density on the metal is zero, can be measured for most metals. It can be said qualitatively that a region of about 0.5 V on either side of this potential is the most susceptible to interference by adsorption of neutral impurities.

When very high solution purity is required, the last stage of purification is often *pre-electrolysis*, employing a large-surface-area electrode, composed of the same metal as the working electrode. The idea behind this procedure is simple: if there is an electroactive impurity in solution, let it be consumed during pre-electrolysis (which is typically conducted for a long time, compared to the duration of the experiment), so that none will be left to interfere with the reaction to be studied in the purified electrolyte. If the impurity is not electroactive, there will be sufficient time for it to be adsorbed on the surface during pre-electrolysis. The electrode used for pre-electrolysis can then be removed from the solution, carrying the impurities with it. One could estimate the time required for pre-electrolysis, but it is probably best to develop the procedure by trial and error, choosing a time beyond which further pre-electrolysis does not help.

We conclude this discussion by stating that maintaining high-purity conditions is essential in measurements of electrode kinetics. Experimental results obtained without proper control of the impurity level cannot be trusted. Yet, the level of impurity required in each case depends on the system being studied and on the method of measurement. One must prudently combine the use of high-purity solvents and chemicals with suitable measuring techniques, to reach the desired level of purity *during measurement* at the lowest cost in materials and effort.

7.2 The Hydrogen Evolution Reaction (HER)

7.2.1 Hydrogen Evolution on Mercury

Mercury electrodes have been studied more than any other type of electrode, because of their ease of purification and the high degree of reproducibility attainable with them. All aspects of hydrogen evolution on mercury have probably been studied at one time or another. On the basis of all experimental

evidence, it is commonly accepted that in this case, the first charge-transfer step is rate-determining, and is followed by fast ion–atom recombination

$$H_3O^+_{soln} + e^-_M \xrightarrow{rds} H_{ads} + H_2O \tag{7.2}$$

$$H_3O^+_{soln} + H_{ads} + e^-_M \xrightarrow{fast} H_2 + H_2O \tag{7.3}$$

The sequence of Eqs. (7.2) and (7.3) represent the Volmer–Heyrovsky reaction route in acid solutions. The coverage by adsorbed hydrogen atoms must be very low, since none has ever been detected, even at the highest overpotentials measured. This also rules out atom–atom recombination as the fast second step, in a Volmer–Tafel reaction route, since the rate of the reaction

$$H_{ads} + H_{ads} \rightarrow H_2 \tag{7.4}$$

is proportional to θ^2, while the rate of step (7.3) is proportional to the first power of θ.

The Tafel slope for this mechanism is $2.3RT/\beta_c F$, and this is one of the few cases offering good evidence that $\alpha \approx \beta$, namely, that the experimentally measured transfer coefficient is equal to the symmetry factor. A plot of log χ versus E, where χ is the dimensionless rate constant, given by

$$\chi = [(12/7)(t/D)]^{1/2} k_{s,h} \tag{7.5}$$

obtained on a dropping mercury electrode in a dilute acid solution, is shown in Figure 7.2. The activation-controlled current density, j_{ac}, corrected for mass-transport limitation, is proportional to the parameter χ. The accuracy shown here is very high and can be achieved only on a hanging dropping mercury electrode (HDME). On solid electrodes, one must accept a significantly

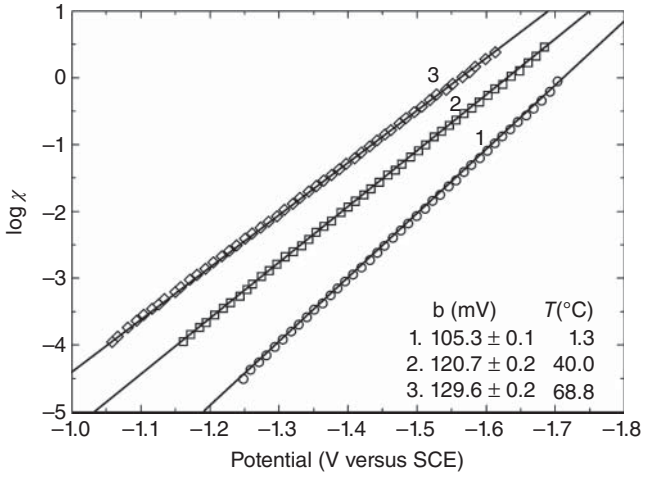


Figure 7.2 Tafel plots for the HER at a hanging dropping mercury electrode (HDME) in 3 mM HCl and 0.8 M KCl. The dimensionless rate constant is given by $\chi = [(12/7)(t/D)]^{1/2} k_{s,h}$. Source: Based on data by E. Kirowa-Eisner, M. Schwarz, M. Rosenblum, and E. Gileadi, *J. Electroanal. Chem*. **381** (1995) 29.

lower level of accuracy and reproducibility. The best values of the symmetry factor obtained in this kind of experiment are close, but not exactly equal to 0.50. It should be noted, however, that the Tafel lines are very straight; that is, β_c is strictly independent of potential over a range of 0.5–0.6 V, corresponding to 4–5 orders of magnitude of current density.

The exchange current density for this system depends on the composition of the solution, but generally it is in the range of 10^{-12}–10^{-10} A cm^{-2}. Mercury is often referred to in the literature, rather loosely, as a "high-overpotential metal." This is a poorly chosen term, since the overpotential clearly depends on the current density and cannot be said to have a specific value for a particular metal. It would be better to describe the situation by saying that mercury is "a low exchange current density metal." On the other hand, if overpotentials are compared at a given current density, then it is higher for mercury and similar "high-overpotential metals" than for platinum and other "low-overpotential metals."

Which metals are similar to mercury in this respect? It turns out that most of the soft metals in groups 5B and 6B of the periodic table (including Pb, Bi, Cd, In, and Sn) behave rather similarly to mercury as cathodes for the HER. It would be presumptuous to claim that we could have predicted this similarity from theory, but upon being confronted with the facts, we can reasonably well explain this result on the basis of the poor catalytic activity of these metals, as shown further.

7.2.2 Hydrogen Evolution on Platinum

It is much more difficult to study the kinetics of the HER on platinum than on mercury. To begin with, the exchange current density is found to be many orders of magnitude higher, in the range of $j_0 = 10^{-4}$–10^{-2} A cm^{-2}, depending on the composition of the solution. It may be recalled that the linear Tafel region, in which the rate of the reverse reaction can be neglected, starts only when $j/j_0 \geq 10$ (corresponding to $\eta/b \geq 1$). Since it is necessary to make measurements over at least two decades of current density in order to obtain a reliable Tafel slope, the measurement on platinum must be extended to rather high current densities, where interference by the uncompensated solution resistance and by mass-transport limitations can be significant. Also, it is found that platinum, being a catalytic metal, adsorbs most impurities very well. Moreover, the potential region in which the HER is studied is close to the potential of zero charge on platinum, where adsorption of neutral species is favored. Thus, a very high level of purification is required in order to obtain reliable data for the HER on platinum electrodes.

The results obtained in acid solutions indicate that there are two distinct mechanisms. At low overpotentials, the atom–atom recombination step (7.4) is believed to be rate-determining. This should yield a Tafel slope of $b = 2.3RT/2F = 29.5$ mV, and a reaction order (at constant potential) of $\rho_1 = 2$, in agreement with experimental results. As the overpotential is increased, the fractional surface coverage θ must also increase. This increases the rate of the atom–atom recombination step, but also that of ion–atom recombination, which occurs in parallel. As θ approaches unity, the rate of step (7.4) can no longer increase, but the rate of the ion–atom recombination step can grow, because it

depends on potential (see Eq. (7.3)). This step then becomes rate-determining. The resulting rate equation (for $\theta \to 1$) is

$$j_2 = nFk_2 c_{b,H_3O^+} \exp\left(-\frac{\beta_c F}{RT}E\right) \tag{7.6}$$

This leads to a Tafel slope of $b = 2.3RT/\beta_c F = 118$ mV for $\beta_c = 0.5$, and a reaction order (at constant potential) of $\rho_1 = 1$. The transfer coefficient α_c is equal to the symmetry factor β_c, as in the case of mercury, but we recall that the Tafel slope is calculated here on the assumption of essentially full coverage, whereas that on mercury was obtained for very low coverage.

Is the assumption of high coverage borne out by experiment? Platinum is a catalytic metal, on which molecular hydrogen is adsorbed spontaneously from the gas phase, dissociating and forming adsorbed atoms on the surface. The surface coverage can be readily measured electrochemically, and indeed it is found to be high at high overpotentials. The difficulty with the foregoing interpretation is that θ is found to be close to unity already at low overpotentials, making it hard to justify the low value of the Tafel slope observed in this region. One way out of this dilemma is to assume that there are two types of adsorbed hydrogen on platinum: a first monolayer, which is already observed at the reversible potential, is partially absorbed and becomes, in a sense, a part of the metal surface. This layer can be removed by anodic oxidation, yielding a measure of the coverage, but it does not participate in the HER at low negative overpotentials. The precursors of the HER are hydrogen atoms adsorbed on top of the first layer at low partial coverage. At high overpotentials, the first layer of adsorbed hydrogen becomes electroactive, and the observed kinetics are as expected on a complete monolayer of adsorbed hydrogen. This interpretation is tentative, in the sense that it has not been confirmed directly. There is, however, indirect evidence that this kind of behavior is possible, and even plausible, from the study of the oxidation of iodide in thin-layer cells. In that case, it was shown experimentally that the first layer of iodide ions adsorbed on the surface of platinum is not electroactive, and iodine is formed from ions in solution, which presumably can be adsorbed on top of the first layer. It was also found that the first adsorbed layer can be oxidized at higher overpotentials, as proposed here for the HER.

The region of micro-polarization, where the j/η plot is linear, can extend to about $\eta/b \leq 0.2$, whereas the linear Tafel region starts at about $\eta/b \geq 1$. As a result, the intermediate region of $0.2 \leq \eta/b \leq 1$ would be left unused, as far as the evaluation of kinetic parameters is concerned. For fast reactions, such as the HER on platinum, this represents a loss of crucial data, since it may be difficult to extend the measurements to overpotentials much above $\eta/b = 1$, because of mass-transport limitations. Fortunately, modern computers allow us to make use of this intermediate region. To do this, we write the full equation for an activation-controlled electrode reaction as follows:

$$j/j_0 = \exp\left(\frac{\alpha_a F}{RT}\eta\right) - \exp\left(-\frac{\alpha_c F}{RT}\eta\right) \tag{7.7}$$

assuming that mass transport has been corrected for. The values of j_0, α_a, and α_c are then calculated by parameter-fitting of the data, according to Eq. (7.7),

over the whole range of overpotentials accessible experimentally. Moreover, for fast reactions, it is often possible to make measurements both at positive and at negative overpotentials, extending even further the range of data from which the kinetic parameters could be obtained.

Metals that are known to be highly catalytic for hydrogenation and dehydrogenation reactions are similar to platinum with respect to the HER. These include palladium, iridium, rhenium, nickel, and cobalt. In the case of palladium, the situation is further complicated by the tendency of atomic hydrogen to diffuse *into* the metal and dissolve in it. Although this effect is most pronounced for palladium, which can absorb hydrogen to an extent represented by the formula $PdH_{0.6}$, hydrogen is also absorbed in iron, nickel, titanium, aluminum, platinum, and some of their alloys. The HER in these cases should be written in the form:

$$H_3O^+_{soln} + e^-_M \rightarrow H_{ads} + H_2O \tag{7.8}$$

The adsorbed hydrogen atom can then follow two parallel paths – either forming molecular hydrogen or being absorbed in the metal. Penetration of hydrogen into the metal is a side reaction for the HER. If it is substantial, as in the case of palladium, the *j/E* plot cannot be analyzed properly, unless the rates of the two parallel reactions are determined as a function of potential and time. This could be performed, for example, by employing a ring disc electrode. Molecular hydrogen would be formed on the disc, and its rate of formation could be determined by oxidizing it back to H_3O^+ on the ring. However, such measurements have rarely, if ever, been reported.

It should be noted here that Pt and some of its alloys are of great importance in the area of fuel cells and water electrolyzers, and in the general context of the so-called "*hydrogen economy*." In these fields, the high value of the exchange current density is a great advantage, and the fact that it hinders our efforts to determine the mechanism is quite irrelevant. Apart of being an excellent catalyst for oxidation of molecular hydrogen (as well as some prospective organic fuels, such as methanol), its stability in acid solutions under usual operation conditions is an important asset. Great efforts have been made to reduce the cost of Pt employed as anodes in fuel cells, by using small amounts of the metal in highly dispersed form, usually on large-surface-area carbon. This is discussed in detail in Chapter 20.

7.3 Possible Paths for the Oxygen Evolution Reaction

Oxygen evolution is a more complex reaction than hydrogen evolution, involving the transfer of four electrons. The overall reaction in alkaline solutions is

$$4(OH)^- \rightarrow O_2 + 4e^-_M + 2H_2O \tag{7.9}$$

whereas, in neutral or acid solutions it is

$$6H_2O \rightarrow O_2 + 4H_3O^+ + 4e^-_M \tag{7.10}$$

Let us write a mechanism (one of many possible) for the reaction shown in Eq. (7.9), assuming that the first step is rate-determining:

$$OH^- \xrightarrow{k_1} OH_{ads} + e_M^- \tag{7.11}$$

This could be followed by a further charge-transfer step, such as

$$OH^- + OH_{ads} \xrightarrow{k_2} O_{ads} + e_M^- + H_2O \tag{7.12}$$

Followed, for example, by atom–atom recombination, to yield

$$2O_{ads} \xrightarrow{k_3} O_2 \tag{7.13}$$

Let us evaluate the kinetic parameters corresponding to this reaction sequence. For the first charge transfer as the rate-determining step, we already know the result, since it is equivalent to step (7.2) for hydrogen evolution on mercury, except for the change in sign. If the second step is assumed to be rate-determining, we also know the result (see Eqs. (6.18)–(6.20)). The Tafel slope changes from $b_a = [2.3RT/(1+\beta_a)F]$ to $b_a = 2.3RT/\beta_a F$ as the partial coverage increases. The reaction order at constant potential, ρ_1, changes from 2 to 1, as θ approaches unity.

Now, let us consider the third case, assuming the atom–atom recombination, Eq. (7.13), to be rate-determining. In this case, the two preceding steps are at quasi-equilibrium. The corresponding equations are

$$\left[\frac{\theta_{OH} + \theta_O}{1 - (\theta_{OH} + \theta_O)}\right] = K_1 c_{OH^-} \exp\left(\frac{F}{RT}E\right) \tag{7.14}$$

$$\left[\frac{\theta_O}{\theta_{OH}}\right] = K_2 c_{OH^-} \exp\left(\frac{F}{RT}E\right) \tag{7.15}$$

Hence, for the rate-determining step one has

$$j_3 = nFk_3\theta_O^2 = nFk_3K_2^2\theta_{OH^-}^2(c_{b,OH^-})^2 \exp\left(\frac{2F}{RT}E\right) \tag{7.16}$$

At low values of the total coverage, $[1 - (\theta_O + \theta_{OH})] \approx 1$, which leads to

$$j_3 = nFk_3K_1^2K_2^2(c_{b,OH^-})^4 \exp\left(\frac{4F}{RT}E\right) \tag{7.17}$$

For this mechanism, the transfer coefficient α_a is 4, the Tafel slope b_a is 15 mV, and the reaction order at constant potential, ρ_1, is 4.

One could write many other pathways and rate-determining steps, and calculate the kinetic parameters for each, following the same line of reasoning. It is important to note that, in a complex reaction sequence, there can be more than one type of adsorbed intermediate on the surface, and some steps may involve the transformation of one kind of adsorbed species to another, either by an electrochemical or by a chemical route.

Two aspects of the oxygen evolution reaction (OER) are common to all electrodes studied so far: (i) the exchange current density is low, of the order of 10^{-10} A cm^{-2} or less, and (ii) a reversible oxygen electrode operating at or near room temperature has not yet been found. At sufficiently high temperatures

(say, in molten salts at about 600 °C, or with high-temperature solid electrolytes operating at around 800 °C or above), the kinetics of the reaction can be sufficiently accelerated to make reversible oxygen electrodes operate as well as reversible hydrogen electrodes do at room temperature.

An important point that is often ignored, or at least not included explicitly in the interpretation of the experimental findings, is that oxygen evolution never occurs on the bare metal surface. By the time the reversible potential for oxygen evolution is reached, an oxide layer has been formed on all metals. At more anodic potentials, where measurements can actually be conducted (remember that j_0 is small, and it takes a high overpotential to drive the reaction at a measurable rate), the oxide film may be several molecular layers thick. On the other hand, oxygen reduction may well occur on the bare metal surface, or one that is covered only by a fraction of a monolayer. Thus, even when Pt is used as the working electrode, it is well to remember that the oxidation and reduction of oxygen on the same metal occur at different surfaces, and may therefore follow entirely different pathways.

Oxides on metals can be divided into three groups: those having a high electrical conductivity, such as RuO_2 and the oxides formed on platinum; those that are semiconductors, such as NiO, TiO_2, the oxides formed on W and Mo; and those that are insulators, such as Al_2O_3 and the oxides formed on Ta and Nb. Oxygen evolution can occur readily on the electrically conducting oxides, and these are the best catalysts for it. On semiconducting oxides, the reaction can still occur, but it may be associated with pitting on the one hand, and further buildup of the oxide layer on the other, causing poor reproducibility and making the interpretation rather dubious. On valve metals, oxide formation is the main reaction occurring during anodic polarization, and the current either decays to zero with time or reaches a constant value, at which the rate of dissolution of the oxide is equal to the rate of its electrochemical formation.

On noble metals (e.g. Pt, Ir, Au), the region of potential in which the OER is studied is far removed from the potential of zero charge in the positive direction. Thus, there is little danger of adsorption of neutral or positively charged impurities, but negatively charged impurities will be readily adsorbed. One may expect, therefore, that the nature of the anion in the electrolyte will influence the measurements substantially, while the type of cation will have little effect. Also, most organic molecules are rapidly oxidized at these potentials, so that the requirements for solution purification are much less severe than in the case of hydrogen evolution.

The preceding statement does not hold true for the oxygen *reduction* reaction (ORR). Using platinum as an example, we note that oxygen evolution is typically studied in the range of 1.5–2.0 V, versus a reversible hydrogen electrode (RHE) in the same solution, while oxygen reduction is studied in the range of 1.0–0.4 V on the same scale. In most of the latter range, the surface is free of oxide (and very sensitive to impurities) if approached from low potentials, whereas it may be covered partially with oxide (and less sensitive to impurities) when approached from higher potentials. This is due to the high degree of irreversibility of formation and removal of the oxide layer on most noble metals.

A discussion of oxygen evolution and, even more, of oxygen reduction, cannot be complete without mentioning the efforts that have gone into improving

the catalytic activity of metals by the use of metal-organic catalysts of the phthalocyanine group. These are large planar molecules, with a metal atom in the center, resembling the structure of the porphyrin molecule, which is involved in the breathing processes of living organisms. The reasoning behind these experiments is to try to imitate nature. If molecular oxygen can be reduced efficiently in living organisms at or near ambient temperatures, perhaps the molecule involved in this reaction, or other metal-organic compounds having a similar structure, could also act as good catalysts for oxygen reduction occurring *in vitro*, in an electrochemical device.

Unfortunately, the analogy is rather poor. Nature operates in an intricate manner, and the porphyrin molecule is only one part of a complex system, which includes various enzymes and other protein molecules. Nevertheless, some progress has been made along these lines, and phthalocyanines and similar molecules may eventually be developed to become potent catalysts for oxygen evolution and reduction.

7.4 The Role and Stability of Adsorbed Intermediates

Intermediates are commonly formed in chemical reactions, as well as in electrode reactions. The preferred mechanism is that which involves the most stable intermediates, since this is the path of lowest Gibbs free energy of activation. For reactions taking place in the gas phase or in the bulk of the solution, the stability of different species can be calculated, or at least estimated, from existing thermodynamic data. This is not the case for electrode reactions. For the HER discussed earlier, a hydrogen atom was assumed to be an intermediate. The standard potential for the formation of this species *in solution*, that is, for the reaction

$$H_3O^+_{soln} + e^-_M \rightarrow H^0_{soln} + H_2O \tag{7.18}$$

is about −2.1 V versus SHE. Thus, this species could not be formed in the potential range over which the HER is studied, even on mercury electrodes. The real reaction we are looking at, however, is not represented by Eq. (7.18), but rather by Eq. (7.2):

$$H_3O^+_{soln} + e^-_M \rightarrow H^0_{ads} + H_2O \tag{7.2}$$

The hydrogen atom formed as an intermediate is stabilized by adsorption. To clarify this point, we should perhaps rewrite Eq. (7.2) in explicit form, as follows:

$$H_3O^+_{soln} + e^-_M + Pt \rightarrow Pt\text{-}H + H_2O \tag{7.19}$$

implying that a chemical bond is formed between the hydrogen atom and a metal atom on the surface. The reversible potential for this reaction is clearly different from that for Reaction (7.2). The difference between them depends on the Pt—H bond energy on the surface. This should not be confused with the bond energy of bulk platinum hydride, or with the energy of formation of a Pt—H species in the gas phase, since platinum atoms on the surface are energetically different from atoms in the bulk, which, in turn, are different from isolated atoms in the

gas phase. It is found that the standard reversible potential for Reaction (7.19) is $E^0 \approx +0.2$ V versus RHE. For the HER:

$$2H_3O^+ + 2e_M^- \rightarrow H_2 + 2H_2O \tag{7.20}$$

the standard reversible potential on the same scale is, of course, zero by definition. We note that the adsorption energy of hydrogen on platinum is so high that it is easier (i.e. it requires a less negative potential) to form the adsorbed species Pt—H, than to form H_2 molecules in solution. In comparison, the Hg—H bond is very weak, and consequently a hydrogen atom is not stabilized on the surface of mercury. The reversible potential for Reaction (7.19), with platinum replaced by mercury, would be close to that for Reaction (7.18). This is why θ_H on mercury cathodes is below the detection limit, even at the highest overpotentials studied. The situation is similar in the case of oxygen evolution, where $(OH)_{ads}$ and O_{ads} are postulated as adsorbed intermediates, even though these radicals are very unstable in the bulk of the solution.

7.5 Adsorption Energy and Catalytic Activity

When one tries to correlate the electrocatalytic activity of metals with some fundamental property of the system, the result is often a "volcano-type" plot, as shown in Figure 7.3.

Such behavior is easy to understand qualitatively, although a quantitative relationship may be hard to derive. The role of a *heterogeneous catalyst* is to adsorb

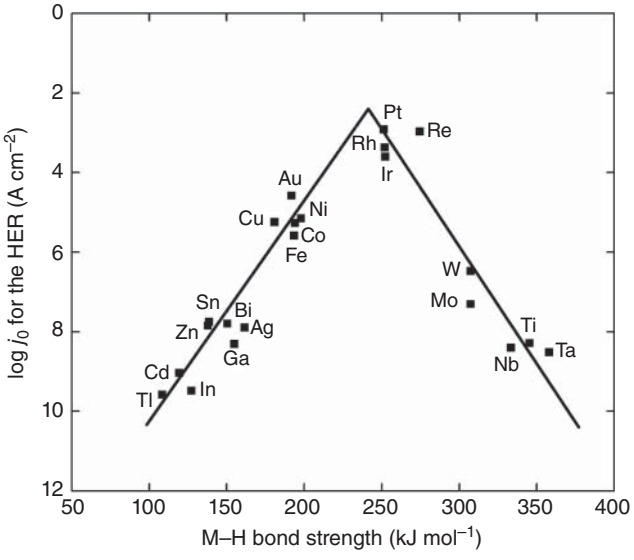


Figure 7.3 The exchange current density for the hydrogen evolution reaction as a function of the M-H bond energy. Source: Reprinted with permission from S. Trasatti, *J. Electroanal. Chem.* **39** (1972) 163.

the reactant or intermediate and transform it to a species that can undergo the desired chemical reaction more readily. If the bonding energy for adsorption is very low, the extent of adsorption will be very small. Moreover, the adsorbed molecules will be weakly bound to the surface and will not be affected by it. As the bonding energy increases, the fractional surface coverage also increases, and the adsorbed species can be modified and activated by its bond to the surface. But, this can be overdone. Beyond a certain point, the coverage approaches saturation, and the rate of reaction can no longer increase with increasing energy of adsorption. Also, if the bonding energy to the surface is too high, the adsorbed intermediate or its product may stick to the surface, effectively poisoning it. In short, the energy of adsorption must be high enough to attach the reactant to the surface, yet low enough to allow it to react there and release the product to the solution. This leads to the so-called "volcano-type" relationship, where the best catalyst is one giving rise to an intermediate value of the bonding energy for adsorption.

A word of caution is appropriate here, in relation to making such correlations. The proper comparison would be one in which the energy of adsorption was changed, while all other properties of the metal were kept constant. This cannot be done, of course. When we compare, say, the rate of hydrogen evolution on gold and on platinum, the heat of adsorption of hydrogen for the two metals is different, *but so is every other property*: the electronic work function, the heat of sublimation of the metal, the potential of zero charge, the free-electron density, the crystal structure, and almost any other property one may think of. Often, some of these properties are correlated (e.g. the work function, the potential of zero charge, and the heat of adsorption are nearly linearly related). Finding a correlation between any two quantities does not necessarily prove that one is caused by, or is directly related to, the other.

The data given on the right-hand side of Figure 7.3, where the exchange current density declines with increasing M—H bond strength, may be questioned, because the metals involved all have an oxide layer that is very hard to remove, so that the rate of hydrogen evolution may have been measured on the oxide rather than on the bare metal, in which case, the M—H bond strength may not be relevant.

Further Reading

1 Newman, J. and Thomas-Alyea, K.E. (2004). *Electrochemical Systems*, 3rd edn. New Jersey, USA: Wiley.
2 Kirowa-Eisner, E., Schwarz, M., Rosenblum, M., and Gileadi, E. (1995). Temperature dependence of the transfer coefficient for the hydrogen evolution reaction on the DME. *J. Electroanal. Chem.* 381: 29–37.
3 Gileadi, E. (1993). *Electrode Kinetics – For Chemists, Chemical Engineers and Materials Scientists*. Wiley-VCH.
4 Trasatti, S. (1972). Work function, electronegativity, and electrochemical behaviour of metals: III. Electrolytic hydrogen evolution in acid solutions. *J. Electroanal. Chem.* 39: 163–184.

8

The Electrical Double Layer (EDL)

8.1 Models of Structure of the EDL

8.1.1 Phenomenology

In Chapter 1 we pointed out that the impedance of the metal/solution interphase is partially capacitive. According to Gauss, the electrical charge of a conductor resides entirely on its surface, and the electric field inside the conductor is zero. In simple cases, the equivalent circuit is that shown in Figure 8.1a. This equivalent circuit overlooks, for example, diffusion effects. The *double-layer capacitance* C_{dl} and the faradaic resistance R_F are inherent properties of the interphase, which we measure experimentally and interpret theoretically. The value of C_{dl}, for example, depends on the metal electrode, the solution composition, and the applied potential. The solution resistance R_S is not a property of the interphase. It can be viewed as an "error term" arising from the fact that the potential in solution is always measured far from the interphase on the molecular scale, typically at a distance of 10^6–10^7 nm (1.0–10 mm).

If the interphase is *ideally polarizable*, the faradaic resistance approaches infinity, and the equivalent circuit shown in Figure 8.1a can be simplified to that shown in Figure 8.1b. The two adjacent phases cannot exchange charges, and the system then behaves like a capacitor, which can be charged by applying an external potential. If it is *ideally nonpolarizable*, the faradaic resistance tends to zero, and the equivalent circuit shown in Figure 8.1c results. In this case, the two phases can exchange charge carriers, ions or electrons, and in the stationary case the potential difference is determined by the difference of the chemical potential of these carriers in the two phases. Real systems never behave ideally, of course; they may approach one extreme behavior or the other, or be anywhere in between. It is also important to remember that both C_{dl} and R_S depend on potential and should be defined in differential form as follows:

$$C_{dl} \equiv \left(\frac{\partial q_M}{\partial E}\right)_{\mu} \quad \text{and} \quad R_F \equiv \left(\frac{\partial E}{\partial j}\right)_{\mu} \tag{8.1}$$

where q_M is the excess charge density on the metal surface. The double-layer capacitance can be measured in a number of ways, some of which will be discussed in this chapter.

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

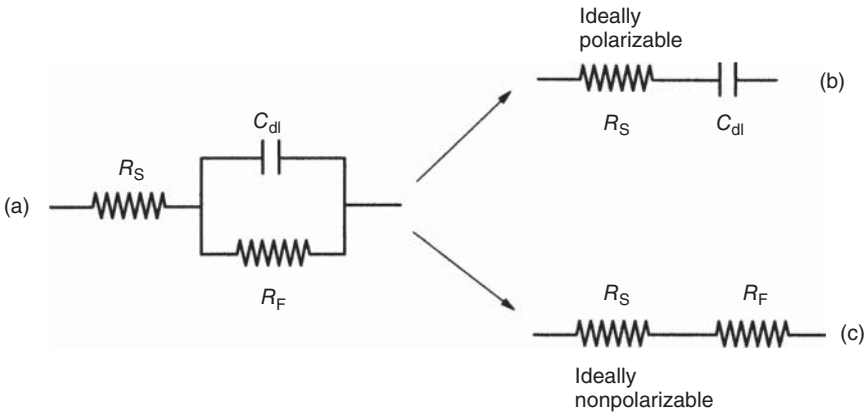


Figure 8.1 The equivalent circuit for an activation-controlled process, showing the three basic circuit elements: the double layer capacitance, C_{dl}, the faradaic resistance, R_F, and the uncompensated solution resistance, R_S. (a) General; (b) ideally polarizable; (c) ideally nonpolarizable.

The system that has been most widely studied is mercury. In addition to being highly reproducible and easy to purify, the interphase it forms in solution is almost ideally polarizable over a relatively wide range of potentials, making both experimental measurement and theoretical interpretation easier. The best results on solid electrodes have been obtained on single-crystal surfaces of gold, which are the most reproducible of solid surfaces. Typical results for mercury are shown in Figure 8.2. The effect of a strongly adsorbed organic molecule on the capacitance–potential curve is also shown. The reader will note that the potential scale is reversed, negative to the right and positive to the left. This

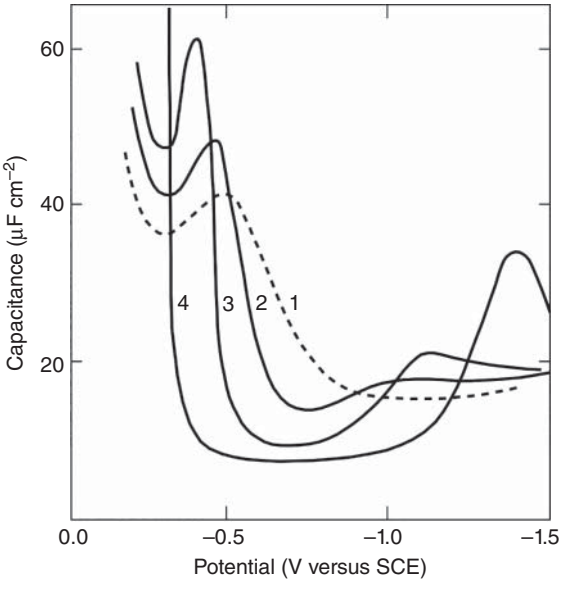


Figure 8.2 Experimental plots of C_{dl} versus E on Hg in 1.0 M KCl, showing the effect of addition of *N*-propylamine concentration: (1) 0, (2) 5×10^{-5} M, (3) 1×10^{-4} M, (4) 5×10^{-4} M. Source: W. Lorenz, F. Möckel, and W. Müller, *Z. Physik. Chem. (N.F.)* **25** (1960) 145.

unfortunate choice was made in the early days of electrochemical research, when the field was dominated by polarography on mercury, and it is still in use in most papers and textbooks when dealing with the mercury electrode. In order to avoid confusion, we have chosen, somewhat reluctantly, to follow the same convention when studies of the double-layer capacitance and of adsorption on mercury are discussed.

One of the most important features seen in Figure 8.2 is the nearly constant value of the capacitance at the far negative end, in the absence of adsorption (curve 1). This value, of about 16 μF cm^{-2}, is essentially independent of the electrolyte used. This observation played an important role in the development of our understanding of the structure of the double layer at the metal/solution interphase, as discussed further.

The other point of great interest is the effect of the organic material. First, we note that the capacitance is much smaller when organic matter is adsorbed. Secondly, we observe that this occurs only in a certain potential range, around the potential of zero charge, E_{pzc} (compare curve 1 with curves 2–4), decreasing both at more negative and at more positive potentials. This result is far from being self-evident, considering that the organic species used here is not charged. Moreover, similar behavior is observed for many neutral organic molecules of widely different structures. These observations form a further cornerstone of our understanding of the structure of the double layer, and in particular, of the factors controlling the adsorption of neutral molecules.

8.1.2 The Parallel-Plate Model of Helmholtz

The first attempt to explain the capacitive nature of the interphase is credited to Hermann von Helmholtz (1879). In his model (see Figure 8.5a below), the interphase is viewed as a parallel-plate capacitor – an excess charge on the surface of the metal attracts an equivalent amount of counter ions to the interphase. The two opposing layers are separated by a certain distance (d in Figure 8.5a), which determines the capacitance. The resulting linear potential drop across the interphase (see Figure 8.5a) is the double-layer potential. It should be noted here that electro-neutrality must be maintained *in the bulk* of all phases, but not at the interphase. Here, there can be an excess charge density on the metal, which we denote q_M, and an excess charge density, q_S, on the solution side of the interphase. The interphase as a whole must be electro-neutral. It follows that for any metal/solution interphase one can write

$$q_M + q_S = 0 \tag{8.2}$$

While the Helmholtz model can explain the existence of a capacitance at the interphase, it cannot explain either its dependence on potential or its numerical value at any potential. The Helmholtz layer specific area differential capacitance (F m^{-2}) for a parallel-plate capacitor, is given by

$$C_H \equiv \frac{\partial \sigma}{\partial V} = \frac{\varepsilon_0 \varepsilon_r}{d} \tag{8.3}$$

where σ is the stored charge density; V is the voltage drop; $\varepsilon_0 \cong 8.854 \times 10^{-12}$ F m^{-1} is the permittivity of vacuum (also known as the permittivity of free space, or the

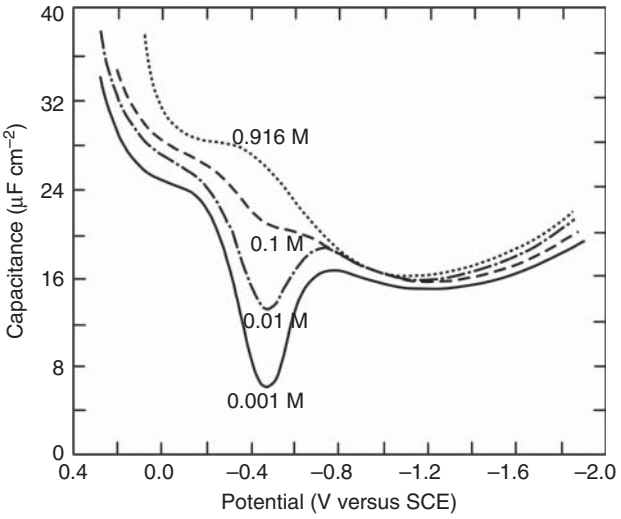


Figure 8.3 Double-layer capacitance on Hg in solutions of different concentrations of NaF at 0 °C. The minima observed in the curves for 0.01 M and 0.001 M NaF show the position of the potential of zero charge, $E_{pzc} = -0.472$ V versus SCE. Source: Data from D.C. Grahame and B.A. Soderberg, ONR *Tech Rep. N°* 14, 1954.

electric constant); and ε_r is the relative permittivity of the medium, i.e. the ratio between the absolute permittivity of the solvent and that of vacuum (dimensionless). For water at 25 °C, $\varepsilon_r = 78.5$. In Eq. (8.3), d is the distance between the two plates (i.e. the thickness of the double layer). Here, it is equal to the radius of the solvated cations. The capacitance calculated for this model, according to Eq. (8.3), should be independent of potential, contrary to the experimental observation shown in Figure 8.3. Moreover, it should be independent of charge and hence, of concentration in solution, while in reality it depends strongly on concentration, as shown in Figures 8.2 and 8.3.

The numerical value obtained from Eq. (8.3), with a typical value of $d = 0.2$ nm, yields a value of 340 μF cm^{-2}. This value is more than 20 times the experimentally observed value of 16 μF cm^{-2} shown in Figure 8.3 at negative potentials, where the solution side of the interphase is dominated by cations.

If a model has failed so utterly in describing the experimental results, how could it withstand the test of time and still merit a mention a century and a half later? A hint may be found by observing the extreme negative end of the plots in Figures 8.2 and 8.3. Over about one volt, from −0.8 to −1.8 V versus SCE, the double-layer capacitance is almost independent of concentration, and only slightly dependent on potential, as expected from the simple parallel-plate model. Part of the above discrepancy can be accounted for by a better choice of the values of ε_r and d in Eq. (8.3), as seen later.

8.1.3 The Diffuse Double Layer Model of Gouy and Chapman

A different approach to interpret the behavior of the double-layer capacitance was taken by Louis Georges Gouy (1909), and later by David Leonard Chapman

(1913). The Gouy–Chapman statistical model describes a rigid charged surface, with a cloud of oppositely charged ions in the solution, the concentration of the oppositely charged ions decreasing with distance from the surface (see Figure 8.5b). This is the so-called *diffuse double layer*. The double layer thickness decreases with increasing valence and concentration. The fundamental premise of this model is that the ions constituting the charge q_S, on the solution side of the interphase, are acted upon by two forces – they interact electrostatically with the excess surface charge on the surface of the metal, q_M, but are also subject to random thermal motion that acts to equalize the concentration throughout the solution.

The equilibrium between these two opposing tendencies is expressed by the well-known *Boltzmann equation*

$$c_i(x) = c_{b,i} \exp\left(-\frac{z_i F}{RT}\phi_x\right) \tag{8.4}$$

where ϕ_x represents the potential at a distance x from the surface of the metal (with respect to the potential in the bulk of the solution, ϕ_S, which is taken as zero), z_i is the charge on the ion, and $c_{b,i}$ is the concentration of the ion in the bulk solution. The concentration of any ionic species at a distance x from the surface, $c_i(x)$, is determined by the exponent of the ratio between the electrostatic energy, $z_i F\phi_x$, and the thermal energy, RT. From here on, the derivation is similar to the better-known derivation of the Debye–Hückel Limiting Law, giving the mean activity coefficients of ions in solution as a function of the ionic strength, which was, however, published about 10 years later.

The potential ϕ_S is related to the charge density per unit volume, $\rho(x)$, by the *Poisson equation*. Since the changes in ϕ_x and $\rho(x)$ are considered only in the direction perpendicular to the surface, this equation takes the simple, one-dimensional, form

$$\left(\frac{\partial^2 \phi_x}{\partial x^2}\right) = -\frac{\rho(x)}{\varepsilon_0 \varepsilon_r} \tag{8.5}$$

The volume-charge density $\rho(x)$ is related to the concentration of all the ions in solution by

$$\rho(x) = F \sum z_i c_i(x) \tag{8.6}$$

One additional equation that is needed to solve this problem is the *Gauss Theorem*, which relates the excess surface-charge density to the gradient of the potential

$$q_M = -\varepsilon_0 \varepsilon_r \left(\frac{\partial \phi_x}{\partial x}\right)_{x=0} \tag{8.7}$$

We shall skip the details of the derivation, which is described in many textbooks, and present the final results, which are valid for symmetrical electrolytes, e.g. NaF or $MgSO_4$ (but not Na_2SO_4 or $MgCl_2$). Note that in Eq. (8.8) the potential ϕ_x is replaced by ϕ_0, since the equations were solved for the potential at the surface, not at some distance x from it:

$$q_M = (8RT\varepsilon_0 \varepsilon_r c_b)^{1/2} \sinh\left(\frac{|z|F}{2RT}\phi_0\right) \tag{8.8}$$

The diffuse-double-layer capacitance is derived from this equation by differentiating q_M with respect to potential, leading to

$$C_{G.C.} = \left(\frac{\partial q_M}{\partial \phi_0}\right)_\mu = |z|F\left(\frac{2\varepsilon_0\varepsilon_r}{RT}c_b\right)^{1/2} \cosh\left(\frac{|z|F}{2RT}\phi_0\right) \tag{8.9}$$

From Eqs. (8.8) and (8.9) it can be seen that, at the potential of zero charge E_{pzc}, where $q_M = 0$, the potential ϕ_0 at the surface also equals zero, and the differential capacitance has its minimum value, which is proportional to the square root of the bulk concentration

$$C_{G.C.}(\min) = |z|F\left(\frac{2\varepsilon_0\varepsilon_r}{RT}c_b\right)^{1/2} \tag{8.10}$$

This is a beautiful model, in that it allows us to calculate the excess surface-charge density and the double-layer capacitance from well-known principles of electrostatics (the Poisson equation, the Gauss theorem, and the Boltzmann statistics equation, coming from thermodynamics). It has, however, one major drawback: it does not predict the correct experimental results! Perhaps it would be more accurate to state that an agreement between the model and the experiment is found only in dilute solutions and over a limited range of potentials, near the potential of zero charge, as seen in Figure 8.4.

Now, we could come up with several reasons to explain why this model would deviate *to some extent* from experiment:

1. In Eq. (8.4) we made the simplifying assumption that the only energy involved in bringing an ion from infinity to a distance x from the surface is the electrostatic energy, $z_i F\phi_x$, neglecting ion–ion interactions, which are bound to be important at higher concentrations.
2. In the derivation of Eq. (8.9) it was tacitly assumed that the relative permittivity, ε_r, is independent of the distance from the electrode. This is only an approximation. Theoretical calculations indicate that ε_r changes from 8 to 10 in the first layer of water molecules at the metal surface to its bulk value of 78 over a distance of 1–2 nm.

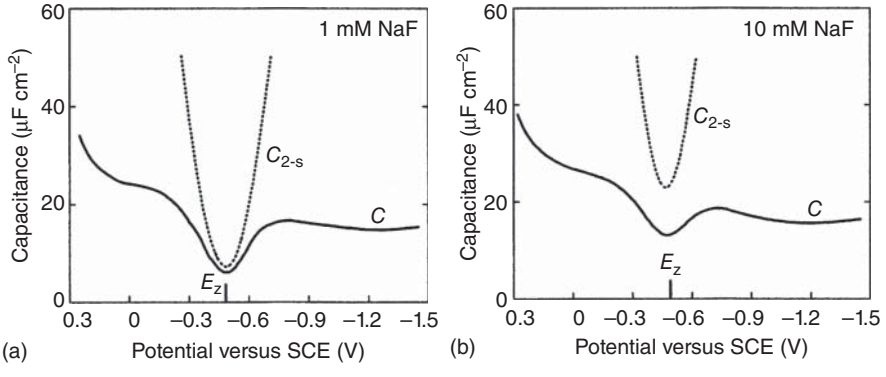


Figure 8.4 A comparison between $C_{G.C.}$, calculated from the Gouy–Chapman model (dashed lines) and experiment (solid lines) for (a) 1 mM NaF, and (b) 10 mM NaF at 25 °C. Source: Based on data from D.C. Grahame, *J. Am. Chem. Soc.* **76** (1954) 4819.

3. Perhaps the largest error is introduced by using the potential ϕ_0 at $x = 0$, which is equivalent to treating the ions as point charges, for which the distance of closest approach to the surface is taken as zero.

All these are valid objections, and the model could be improved, at least in principle, by deriving appropriate correction terms. However, none of these could explain the total disagreement between theory and experiment at a potential of, say, 0.5 V on either side of the potential of zero charge, as well as in concentrated solutions. As in the discussion of the Helmholtz model, we might ask ourselves, why a theory that is in almost total disagreement with experiment is discussed at all! It turns out that a clever combination of the Helmholtz and the Gouy–Chapman models, resulting from a clear physical grasp of the situation at the interphase, yields quite good agreement between theory and experiment, as we shall see next.

8.1.4 The Stern Model

The puzzle was solved by Otto Stern (1924), who showed that good agreement between theory and experiment could be achieved, once it was realized that both the Helmholtz and the Gouy–Chapman models are valid and exist simultaneously (see Figure 8.5c). Stern postulated a layer of ions at the surface, which gives rise to the Helmholtz (compact) part of the double layer, marked as C_H or $C_{M\text{-}2}$. Outside this layer there is an ionic space charge, which constitutes the Gouy–Chapman (diffuse) double layer, marked as $C_{G.C.}$ or $C_{2\text{-}S}$. For electro-neutrality across the interphase, one must still have $q_M + q_S = 0$, but the charge on the solution side is partly in the compact layer and partly in the diffuse layer.

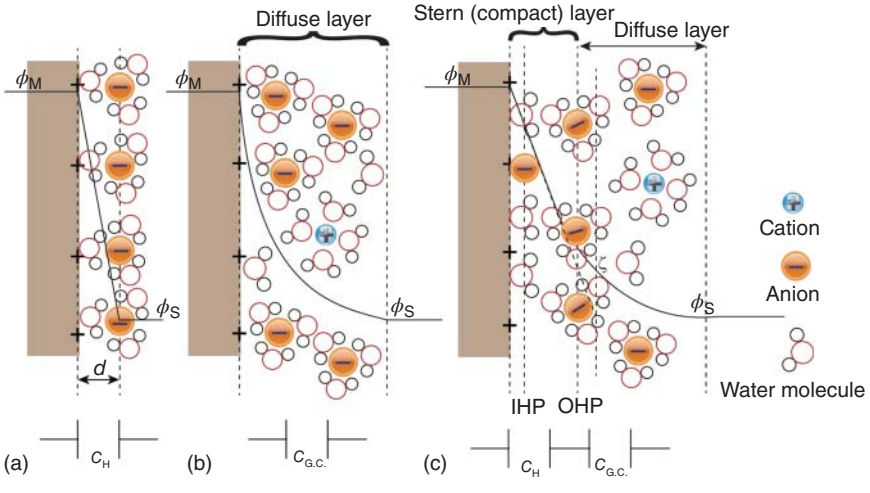


Figure 8.5 Schematic representation of: (a) the Helmholtz model, (b) the Gouy–Chapman model, and (c) the Stern model, for the structure of the electrical double layer in the vicinity of a positively charged electrode surface (anode).

With this model in mind, we can think of the potential drop between the metal and the solution as being divided into two segments:

$$\phi_M - \phi_S = (\phi_M - \phi_2) + (\phi_2 - \phi_S) \tag{8.11}$$

The potential ϕ_2 is called the potential of the *outer Helmholtz plane (OHP)*. The diffuse double layer starts at the outer Helmholtz plane (see Figure 8.5c), where the potential is ϕ_2. It is this value of the potential, rather than ϕ_0, that must be used in Eqs. (8.8) and (8.9), to relate the surface-charge density and the diffuse-double-layer capacitance to the potential. The outer limit of the Helmholtz layer (also known as the Stern layer, or the ion-free layer) comprises the ions that are closest to the electrode surface, but are not specifically adsorbed. They retain their solvation spheres intact, and are bound only by electrostatic forces. If all these ions are equivalent, their centers define the OHP. It should be noted here that the distance of the OHP from the metal surface is determined (in the framework of the Stern model) by the radius of the hydrated cations in solution, and the potential in that plane is ϕ_2. The next refinement of the model takes into account also a layer of water molecules adsorbed on the surface. A discussion of this phenomenon and its effect on the double-layer capacitance is made further.

There is a similar quantity called the *inner Helmholtz plane (IHP)*, where the potential is ϕ_1. This is determined by the radius of the anions in solution, which tend to be specifically adsorbed on the metal surface. If only one type of molecule or ion is adsorbed, then their centers define the IHP (see Figure 8.5c).

Differentiating Eq. (8.11) with respect to q_M (while setting $\phi_S = 0$) yields

$$\left[\frac{\phi_M}{q_M}\right]_\mu = \left[\frac{\phi_M - \phi_2}{q_M}\right]_\mu + \left[\frac{\phi_2}{q_2}\right]_\mu \tag{8.12}$$

This can also be written in the form

$$\frac{1}{C_{dl}} = \frac{1}{C_H} + \frac{1}{C_{G.C.}} \tag{8.13}$$

Here, C_{dl} is the experimentally measured double-layer capacitance. Equation (8.13) has the usual form for two capacitors connected in series, according to the model postulated by Stern.

The important thing to note in Eq. (8.13) is that in a series combination of capacitors, it is always the *smallest* capacitor that will predominantly determine the overall capacitance observed. This can explain qualitatively the observation shown in Figures 8.3 and 8.4. Agreement between the Gouy–Chapman model and experiment is found only in dilute solutions, not exceeding 0.1 M, and only in the vicinity of the potential of zero charge. Since $C_{G.C.}$ increases linearly with $(c_b)^{1/2}$ and is proportional to $\cosh\left(\frac{|z|F}{2RT}\phi_2\right)$, it follows from Eq. (8.13) that it can no longer be observed at high concentration and at potentials far away from the potential of zero charge.

It is evident now why the models of Helmholtz and Gouy–Chapman were not discarded. While each one alone fails completely when compared with experimental data, a series combination of the two yields reasonably good agreement. There is room for improvement and refinement of the model. The

next improvement of the Stern model was made by David Grahame (1947) who proposed that some ionic or uncharged species can penetrate the Stern layer, although the closest approach to the electrode is normally occupied by solvent molecules. This could occur if ions lose their solvation shell as they approach the electrode. He termed ions in direct contact with the electrode "specifically adsorbed ions." This model proposed the existence of three regions: (i) the IHP plane that passes through the centers of the specifically adsorbed ions, (ii) the OHP that passes through the centers of solvated ions, that are not specifically adsorbed, at the distance of their closest approach to the electrode, and (iii) the diffuse layer which is the region beyond the OHP. These models bring theory and experiment close enough for us to believe that they do describe the real situation at the interphase.

The distances related to the IHP and OHP depend on the size of the ions involved. Taking into account the hydration of the surface, the typical thickness of the interface is in the range of 0.5–0.7 nm, ignoring the Gouy–Chapman layer that may extend up to a distance of about 10 nm.

Figure 8.5 summarizes schematically the three main models for the structure of the double layer around a positively charged metal electrode (anode). From Figure 8.5a it is evident that the thickness d of the Helmholtz double layer can be approximated as the radius of the solvated anions, and that all the counter ions are assumed to be adsorbed at the electrode surface. The electrode potential, ϕ_M, declines linearly to the solution potential at the plane of the centers of the counter ions, ϕ_S. The Helmholtz capacitance, C_H, is marked. Figure 8.5b illustrates the excess negative charge within the diffuse layer that compensates the excess positive charge at the metal surface. Note that solvated anions, solvated cations, and free water molecules all exist within the diffuse layer. The Gouy–Chapman capacitance, $C_{G.C.}$, is also marked. Figure 8.5c illustrates the compact layer and the diffuse layer in the Stern model. A layer of water molecules is adsorbed on the metal surface, along with specifically adsorbed anions. The centers of these adsorbed species establish the IHP, while the centers of the first layer of non-adsorbed solvated anions effectively establish the OHP. Within the diffuse layer, the slipping plane is marked by vertical dash line, and its intersection with the potential curve yields the zeta potential, ζ. The experimentally measured double-layer capacitance represents the two capacitors connected in series, as noted in this figure.

We have failed to discuss so far the numerical value of the capacitance of the compact layer C_H and its dependence on potential (or charge), both of which are in disagreement with the simple parallel-plate capacitor model proposed originally by Helmholtz. These issues, and the important effect of the solvent at the interphase, are discussed in the next section.

8.1.5 The Role of the Solvent at the Interphase

The solvent can influence the structure of the interphase in many ways:

1. Different ions in solution are hydrated to different degrees, depending on the size of the ion. The interaction of ions with the surface depends on their degree of hydration. In aqueous solutions, most cations are strongly hydrated. As a

result, their interaction with the surface is mainly electrostatic. In contrast, most anions are not hydrated. As a result, anions can be in direct contact with the surface, allowing specific chemical interactions to play a role. Therefore, their adsorption is called *specific* or *contact* adsorption.

2. The surface itself is hydrated. It can be viewed as a giant ion, having a large number of charges. The field near the surface (inside the compact part of the double layer) is comparable to the field around an ion in the bulk of the solution. Thus, hydration of the electrode surface accompanied by breakdown of the bulk structure of the solvent is expected to occur, just as it does in the inner hydration shells of ions. Since the water molecules are in direct contact with the surface, short-range chemical interactions with the metal electrode cannot be ignored. The chemical part of the adsorption energy of water on mercury is not very large, but it is enough to explain the experimentally observed lack of symmetry of the plots of q_M versus E around the potential of zero charge. It may be much larger on solid electrodes, particularly on catalytic metals such as platinum and nickel.
3. The solvent also acts as a dielectric medium, which determines the electrostatic field $d\phi/dx$ and the energy of interaction between charges. The dielectric constant ε depends on the inherent properties of the molecules (mainly, their permanent dipole moment and polarizability) and on the structure of water as a whole. Water is unique in this sense. It is highly associated in the liquid phase and so has a dielectric constant of 78 (at 25 °C), which is much higher than that expected from the properties of the individual molecules. When it is adsorbed on the surface of an electrode, inside the Helmholtz layer, the structure of bulk water is changed, and the molecules are essentially immobilized by the high electrostatic field. Consequently, the appropriate relative permittivity to be used in this region has been estimated to be about 8–10. Farther out in the solution, the bulk structure of water is rapidly regained. The dielectric constant is estimated to reach its bulk value within about 1–2 nm of the surface. In the discussion of the parallel-plate model of Helmholtz, a capacitance of 340 μF cm^{-2} was calculated using $\varepsilon_r = 78$ and $d = 0.2$ nm. This value is somewhat higher than the capacitance of about 16 μF cm^{-2} observed experimentally at extreme cathodic potentials (with respect to the potential of zero charge, E_{pzc}). If, however, we use $\varepsilon_r = 8$–10 instead, we find agreement with experiment for $d = 0.44$ nm, which is consistent with the Stern model, ignoring the adsorbed layer of water on the surface.

What is the "correct" value of the thickness of the parallel-plate capacitor in Eq. (8.3)? The distance of closest approach of a cation is the sum of the diameter of a water molecule (0.27 nm) and the radius of the hydrated ion, the so-called Stokes radius (calculated from electrolytic conductivity data), which is in the range of 0.1–0.3 nm for most cations. The calculated values of d are in the range of 0.5–0.7 nm, consistent with the experimental results.

The distance of the IHP is always smaller than that of the OHP, since there is no layer of water molecules between the anions and the surface. The absence of such a layer can qualitatively explain why the value of C_H is higher on the positive side

of E_{pzc} than on the negative side, taking the parallel-plate model one step closer to the experimental observations.

From an examination of the C_{dl} versus E curves in Figures 8.2–8.4, it is clear that not all the experimental details have been accounted for. There are further effects resulting from lateral ion–ion, ion–dipole, and dipole–dipole interactions. For example, the "hump" observed on the positive side of E_{pzc} can be explained in terms of these interactions, which we shall not discuss, since they are not essential for a basic understanding of the structure of the double layer.

Which ions are specifically adsorbed? It depends, of course, on the metal, but detailed and accurate data available for mercury show that anions, which are generally not hydrated, tend to be specifically adsorbed. This includes most of the anions, but not F^-. Also, some highly symmetrical anions such as ClO_4^-, ReO_4^-, BF_4^-, and PF_6^- are not specifically adsorbed on mercury. Most cations are also not specifically adsorbed on mercury. Cesium, which was found to be specifically adsorbed to some extent, is an exception. Also, large organic cations of the tetra-alkyl ammonium type (for example, $[(C_2H_5)_4N]^+$) are found to be specifically adsorbed. While this information has been most valuable for the understanding of the double-layer structure, it should be borne in mind that mercury is a rather inert, non-catalytic surface. Thus, the structure of the double layer at the surface of catalytic metals could be quite different, and should be considered for each solid metal separately.

Further Reading

1 Bard, A.J., Inzelt, G., and Scholz, F. (ed.) (2012). *Electrochemical Dictionary*. Berlin Heidelberg: Springer-Verlag.

2 Wang, H. and Pilon, L. (2011). Accurate simulations of electric double layer capacitance of ultramicroelectrodes. *J. Phys. Chem. C* 115: 16711–16719.

3 Sharma, P. and Bhatti, T.S. (2010). A review on electrochemical double-layer capacitors. *Energy Convers. Manage.* 51: 2901–2912.

4 Zhang, G. (2010). Simulating the electrical double layer capacitance. Proc. of the COMSOL Conf., Boston.

5 Guidelli, R. and Schmickler, W. (2000). Recent developments in models for the interface between a metal and an aqueous solution. *Electrochim. Acta* 45: 2317–2338.

6 Attard, P. (1996). Electrolytes and the electric double layer. In: *Advances in Chemical Physics*, vol. 92 (ed. I. Prigogine and S.A. Rice). New Jersey, USA: Wiley.

7 Bockris, J.O.'.M. and Khan, S.U.M. (1993). The interphasial structure. In: *Surface Electrochemistry*, Chapter 2. New York, USA: Springer Science.

8 Lipkowski, J. and Ross, P.N. (ed.) (1993). *Structure of Electrified Interfaces*. New York, USA: Wiley-VCH.

9 Lorenz, W., Möckel, F., and Müller, W. (1960). Zur adsorptionsisotherme organischer moleküle und molekülionen an quecksilberelektroden, I. *Z. Phys. Chem.* 25: 145–160.

10 Grahame, D.C. (1954). Differential capacity of mercury in aqueous sodium fluoride solutions. I. Effect of concentration at 25°. *J. Am. Chem. Soc.* 76: 4819–4823.
11 Grahame, D.C. and Soderberg, B.A. (1954). Thermodynamic properties of the double layer. *Technical Report No. 14*. Office of Naval Research, Amherst College.
12 Grahame, D.C. (1947). The electrical double layer and the theory of electrocapillarity. *Chem. Rev.* 41: 441–501.
13 Stern, O. (1924). Zur theorie der elektrolytischen doppelschicht. *Z. Elektrochem.* 30: 508–516.
14 Helmholtz, H. (1853). Ueber einige gesetze der vertheilung elektrischer Ströme in körperlichen leitern mit anwendung auf die thierisch-elektrischen versuche (in German). *Ann. Phys. Chem.* 165: 211–233.

9

Electrocapillary

9.1 Thermodynamics

9.1.1 Adsorption and Surface Excess

The creation of an interphase is a symmetry-breaking process. Even if the two phases in contact are entirely homogeneous, the molecules inside the interphase experience a different average force in each of the two directions perpendicular to the interphase. As a result, the Gibbs free energy of a molecule in the interphase is different from its value in the bulk phase. This effect is not restricted to a single layer of molecules, and it can extend to a distance of several molecular diameters into the solution. Indeed, it is possible to envision a region on either side of the interphase in which the energy of the molecules is different from that in the two bulk phases. This region is referred to as the *interphase*. When only neutral species are involved, the interphase cannot be more than a few molecular layers thick, since the forces involved are "chemical" and decay quickly with distance (see Figure 2.1). When charged particles (i.e. ions in solution in the metal/electrolyte interphase) are considered, the interphase can extend much further, since the electrostatic interactions between charges decrease linearly with distance. The best example of such behavior is the diffuse double layer discussed in Section 8.1.3. In the discussion of the relevant theory, we want to determine the variation of the potential ϕ_2 at the outer Helmholtz plane (OHP) with the excess surface-charge density q_M. The resulting equation is

$$\phi_x = \phi_2 \exp(-\kappa x) \tag{9.1}$$

where κ is the so-called *reciprocal Debye length,* given by

$$\kappa = \left[\frac{2z^2F^2}{\varepsilon_o \varepsilon \; RT}\right]^{1/2} (c_b)^{1/2} \tag{9.2}$$

Substituting the numerical values in this equation, we find that for a 1 mM solution of a symmetrical 1-1 electrolyte, $\kappa^{-1} = 10$ nm. Looking again at Eq. (9.1), it is noted that at a distance $x = \kappa^{-1}$ from the metal, the potential is

$$\phi_\kappa = \phi_2/e = 0.37\phi_2 \tag{9.3}$$

and this is often taken to be "the thickness" of the diffuse double layer, but should more correctly be referred to as "the characteristic length" of the diffuse double

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

layer. A better choice for defining the thickness might be to use $x = 3\kappa^{-1}$, since at this distance ϕ_2 has decayed to 5% of its value at $x = \kappa^{-1}$, but this does not matter, from the conceptual point of view.

The thickness of the diffuse double layer is seen to be a function of concentration. In a 1.0 M solution, $\kappa^{-1} = 0.3$ nm, which is of the order of magnitude of the thickness of the OHP. This shows that the diffuse double layer no longer exists in 1.0 M solutions.

The asymmetry at the surface can affect the energy of different species in solution to different degrees. If, for example, equal amounts of methanol and *n*-butanol are dissolved in water, the concentration of *n*-butanol in the interphase will be higher. This comes about because the transfer of *n*-butanol from the bulk to the interphase decreases the total Gibbs energy of the system more than does the transfer of methanol.

It is necessary, at this point, to define a new quantity, called the *surface excess*, Γ, such that

$$\Gamma \equiv \int_0^\infty (c - c_\mathrm{b})\mathrm{d}x \tag{9.4}$$

The surface excess, Γ, represents the total amount of the relevant species in a cylinder of unit cross-section, extending from the surface of the metal (i.e. at $\phi_x = 0$) into the bulk of the solution, less the amount that *would have been* in the same volume, had there been no interphase. The definition of the surface excess is shown graphically in Figure 9.1. Fortunately, we do not have to determine how far exactly the interphase extends. The function $(c - c_\mathrm{b})$ is integrated to "infinity" – far enough into the bulk so that its value will have become negligible. It should be noted that, on the scale of interest for interphases, "infinity" is not very far. In fact, it is less than 1 μm in the direction perpendicular to the interphase! For example, for $\kappa^{-1} = 10$ nm and a distance of 1.0 μm from the surface, one has, according to Eq. (9.1)

$$\phi_x = \phi_2 \exp(-100) = 3.7 \times 10^{-44} \tag{9.5}$$

We recall, however, that every potential must be referred to some (arbitrarily chosen) zero. When considering the metal/solution interphase in electrochemistry, one usually chooses $\phi_\mathrm{S} = 0$. Thus, Eq. (9.5) could be written as

$$(\phi_x - \phi_\mathrm{s}) = (\phi_2 - \phi_\mathrm{s}) \exp(-100) = 3.7 \times 10^{-44} \tag{9.6}$$

It follows that, at a distance of 1.0 μm, the potential caused by the diffuse double layer is exactly equal to that in the bulk of the solution, for all purposes. Infinity is less than 1.0 μm from the interphase!

The surface excess is an integral quantity. This has the advantage of relieving us from the need to define the boundary of the interphase. On the other hand, measurement of Γ cannot yield any information concerning the variation of the concentration *inside* the interphase, although this can be calculated on the basis of a suitable model. Another point to remember is that the surface excess, as defined here, can have either positive or negative values, as seen in Figure 9.1.

In the case of adsorption of an intermediate in electrode kinetics, we have discussed the extent of adsorption in terms of the fractional surface coverage, θ,

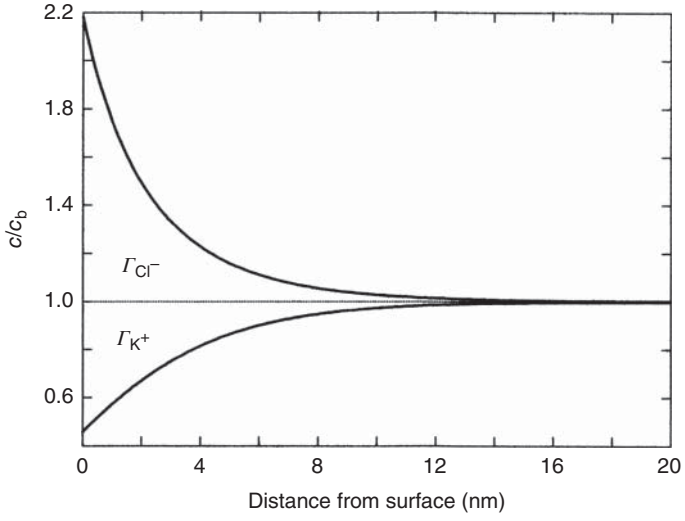


Figure 9.1 Illustration of the concept of the surface excess, Γ. The lines represent the dimensionless concentrations of positive and negative ions, as a function of distance from the mercury surface, calculated for 0.01 M KCl, $q_M = +10\ \mu C\ cm^{-2}$, in the absence of specific adsorption.

which was defined as

$$\theta \equiv \frac{\Gamma}{\Gamma_{max}} \tag{6.9}$$

In this equation, Γ is the surface concentration of the adsorbed species, and Γ_{max} is the value of Γ needed to form a complete monolayer. We recall that the intermediates discussed (such as a neutral H_{ads} atom or an OH radical) are unstable in solution, and can be formed only on the surface. In this case, the surface concentration is equal to the surface excess, and it is appropriate to use the same symbol for both. It should be noted though that Eq. (9.4) does not show a clear value of Γ_{max} and, in fact, Γ could exceed the amount corresponding to a monolayer, because the material is distributed over a thickness that can correspond to several molecular layers.

9.1.2 The Gibbs Adsorption Isotherm

The most fundamental equation governing the properties of interphases is *the Gibbs adsorption isotherm*

$$-\mathrm{d}\sigma = \sum \Gamma_i \,\mathrm{d}\mu_i \tag{9.7}$$

The *surface tension*, σ, is given in units of force per unit length ($1\ N\ m^{-1} = 10^3$ dyn cm^{-1}). It is related to the *two-dimensional surface pressure*, Π ($1\ J\ m^{-2} = 10^3$ erg cm^{-2}), by the simple equation

$$-\mathrm{d}\sigma = d\Pi \tag{9.8}$$

It is also the *excess surface Gibbs energy* per unit area, namely the extra Gibbs free energy added to a system as a result of formation of the interphase. When we speak of surface tension, we have a mechanical model in mind, as though there was an imaginary membrane on the surface, pulling it together with a force (which can readily be measured using a Langmuir–Blodgett trough). The notion of an excess surface Gibbs energy is, of course, purely thermodynamic.

We noted earlier that the driving force in chemistry is the decrease of Gibbs free energy. Thus, a system will change spontaneously in the direction of decreasing surface tension. This leads to two observations:

(i) A pure phase always tends to assume a shape that creates the minimum surface area per unit volume. This is why droplets of a liquid are almost spherical (they are completely spherical in the absence of gravity, in a spacecraft orbiting the Earth, for example).
(ii) When a solution is in contact with another phase, the composition of the interphase differs from that of the bulk in such a manner as to minimize the total excess surface Gibbs free energy of the system.

The second observation represents the essence of the physical meaning of the Gibbs adsorption isotherm. The adsorption of any species in the interphase $\Gamma_I > 0$ must always cause a decrease in the Gibbs free energy of the surface ($d\gamma < 0$), since it is the reduction in this Gibbs free energy that acts as the driving force for adsorption to occur.

The Gibbs adsorption isotherm is derived in many standard textbooks of physical chemistry and surface chemistry. We shall not repeat this derivation here. Rather, we show how this isotherm is modified when it is applied in electrochemistry.

9.1.3 The Electrocapillary Equation

The surface tension of an electrode in contact with an electrolyte depends on the metal–solution potential difference, ${}^{M}\Delta^{S}\phi$. The equation describing this dependence is called the *electrocapillary equation*. It follows by simple logic from the Gibbs adsorption isotherm. Thus, the sum $\sum \Gamma_i \, d\mu_i$ in Eq. (9.7) should represent the surface excess (or deficiency, i.e. negative surface excess) of all the species in the interphase. On the solution side, there are terms of the type ($\Gamma_{Cl^-} \, d\mu_{Cl^-}$) and ($\Gamma_{RH} \, d\mu_{RH}$) for charged and neutral species, respectively, where the subscript "RH" stands for an unspecified organic molecule. On the metal side, the surface excess is expressed by q_M, and the chemical potential is replaced by the electrical potential, E, measured with respect to a chosen reference electrode. This potential differs from the metal–solution potential difference only by a constant.

A general form of the electrocapillary equation can, hence, be written as

$$-d\sigma = q_M dE + \sum \Gamma_i \, d\mu_i \tag{9.9}$$

Several equations follow directly from this equation. The partial derivative of σ with respect to potential, at constant composition of the solution, yields the excess surface charge density

$$-(\partial\sigma/\partial E)_{\mu_i} = q_M \tag{9.10}$$

and the second derivative with respect to potential yields the double-layer capacitance

$$-(\partial^2\sigma/\partial E^2)_{\mu_i} = (\partial q_M/\partial E)_{\mu_i} = C_{dl} \tag{9.11}$$

These equations relate the surface tension, or excess surface Gibbs free energy, to two very important electrical characteristics of the interphase: the charge density q_M and the double-layer capacitance, C_{dl}. It should be noted that these are purely thermodynamic relationships, not based on any model. The only assumption made in the derivation of Eqs. (9.9)–(9.11) is that the interphase is ideally polarizable, that is, charge cannot cross the interphase.

It follows from Eq. (9.10) that the maximum of the electrocapillary curve is given by $q_M = 0$. In other words, the potential of zero charge, E_{pzc}, coincides with the potential of the electrocapillary maximum. The double-layer capacitance can be obtained by double differentiation of the surface tension with respect to potential, and the surface tension can be obtained by double integration of the dependence of C_{dl} on E. The situation is not entirely symmetrical, however. For double differentiation, all one needs is very accurate data of σ versus E. For double integration, one also needs two constants of integration. These are the coordinates of the electrocapillary maximum, namely E_{pzc} and σ_{max}. For liquid electrodes (e.g. mercury and some amalgams), both can be readily measured with high accuracy. For solid electrodes, it is possible to measure E_{pzc}, but reliable values of the surface tension as a function of potential are hard to come by. Thus, for solid electrodes, one can integrate the double-layer capacitance data to obtain q_M as a function of potential, but the second integration, needed to obtain the electrocapillary curve, cannot be readily performed.

The third important relationship that follows from the electrocapillary equation is

$$\Gamma_i = -(\partial\sigma/\partial\mu_i)_{E,\mu_{j\neq i}} = -\frac{1}{2.3RT}(\partial\sigma/\partial \log a_i)_{E,\mu_{j\neq i}} \tag{9.12}$$

This equation can be used to determine the surface excess (in effect, the extent of adsorption in the interphase) of any one species, from the variation of the surface tension with the activity of this species in solution, maintaining the activities of all other species, as well as the potential, constant.

Consider now the application of Eq. (9.9) to a specific system. We choose for this presentation the simple case of a mercury electrode in contact with an aqueous solution of KCl, and a Ag/AgCl reference electrode. This cell can be represented by

$$Cu' \| Ag \| AgCl \| KCl, H_2O \| Hg \| Cu''$$

where the double vertical lines represent separation between different phases (which are in physical contact with each other) and Cu′ and Cu″ are copper wires connected to the terminals of a suitable source of variable potential. Sparing the reader the tedium of going through a long derivation, we arrive at the final equation, which is

$$-d\sigma = q_M\, dE_- + \Gamma'_{K^+}\, d\mu_{KCl} \tag{9.13}$$

where the new symbol, $E\_$, was introduced to show that this specific equation applies to the case in which the reference electrode is reversible with respect to the anion in solution. Had we used a reference electrode that is reversible with respect to the positive ion in solution, Eq. (9.13) would have been replaced by

$$-\mathrm{d}\sigma = q_{\mathrm{M}}\,\mathrm{d}E_{+} + \Gamma'_{\mathrm{Cl}^-}\,\mathrm{d}\mu_{\mathrm{KCl}} \tag{9.14}$$

The symbols Γ'_{K^+} and Γ'_{Cl^-} stand for the *relative surface excess* of the marked species, defined as

$$\Gamma'_{\mathrm{K}^+} \equiv \Gamma_{\mathrm{K}^+} - \Gamma_{\mathrm{H_2O}}\left(\frac{X_{\mathrm{KCl}}}{X_{\mathrm{H_2O}}}\right) \tag{9.15}$$

The need to introduce this function arises from the fact that the derivative given by Eq. (9.11) represents an experiment that is, strictly speaking, impossible to perform. Thus, in a system containing n components, it is impossible to change the concentration of only one of these components, leaving *all other concentrations* constant. However, from the definition of Γ'_{K^+} it follows that it is very close to Γ_{K^+} in a fairly dilute solution, and this refinement can usually be ignored (for example, in a 0.1 M solution of KCl the ratio of mole fractions in Eq. (9.15) is about 0.002). It must, however, be taken into account in a rigorous treatment of the problem, particularly in more concentrated solutions, or when mixed solvents of comparable concentrations are employed.

9.2 Methods of Measurement and Some Results

9.2.1 The Electrocapillary Electrometer

The measurement of surface tension is an old trade in science. There are consequently many methods of determining this quantity. In chemistry, the surface tension is measured as a function of the solvent, the composition of the solution, and the nature of the two phases. In electrochemistry, the potential is an added variable. The *electrocapillary curve* is a plot of the surface tension of a liquid–metal electrode (usually mercury) versus potential, at a given composition of the solution. This type of measurement is then repeated in solutions of different composition, to obtain the surface excess of the appropriate species, employing Eq. (9.12).

To be more exact, we should be talking about the *interphacial tension*, which is the surface tension between two specified phases. In electrochemistry it is customary to use the term *surface tension* to refer to the *interphacial* tension at the metal/solution interphase, or more generally, at the interphase between an electronic and an ionic conductor.

We recall from our elementary science classes the phenomenon of *capillary rise*, shown in Figure 9.2a. When a series of glass capillaries is inserted into water, the water will rise in the capillaries to a height that is inversely proportional to the internal radius of the capillary. The equation describing this behavior is the *Young–Laplace equation*, which can be written as

$$\Delta p = \frac{2\sigma}{r} \tag{9.16}$$

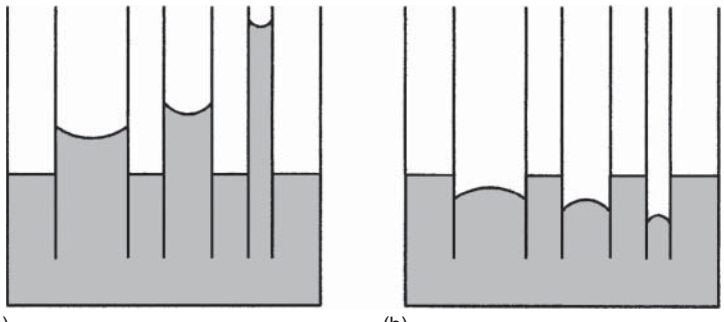


Figure 9.2 (a) Capillary rise observed for a wetting liquid (water), and (b) capillary depression observed for a non-wetting liquid (Hg). Note that this figure is not drawn to scale.

This equation shows the pressure difference Δp needed to keep two phases in mechanical equilibrium, if the interphase has a radius of curvature r. One should distinguish here between *thermodynamic* equilibrium and *mechanical* equilibrium. The former is defined by the condition of minimum Gibbs free energy of the system. The latter represents the condition that the vector sum of all forces acting on the interphase is zero. When the interphase is flat, r goes to infinity and $\Delta p \to 0$. In the present example, the pressure in each capillary is related to the height of the liquid, since

$$\Delta p = \frac{\text{force}}{\text{area}} = \frac{\pi r^2 h \rho g}{\pi r^2} = h \rho g \tag{9.17}$$

where ρ is the density of the liquid and g is the standard acceleration of gravity. The height of the liquid in the capillary is therefore given by

$$h = \frac{\Delta p}{\rho g} = \frac{2\sigma}{r} \frac{1}{\rho g} \tag{9.18}$$

Thus, the capillary rise is determined by the surface tension and the density of the liquid on the one hand, and by the radius of curvature, which in this case equals the inner radius of the capillary, on the other hand.

If we replace water by mercury, a *capillary depression* will be observed, instead of a capillary rise, as shown in Figure 9.2b. The simple physics behind this is that water wets glass, whereas mercury does not. In everyday language, we could say that there is an affinity between water and glass, causing the liquid to "crawl up" the capillary. To force the water in the capillary back to its level in the vessel, it would be necessary to apply a pressure (given by the Young–Laplace equation) that is inversely proportional to the radius of the capillary.

There is no affinity between mercury and glass. Thus, mercury must be forced to enter the capillary. One would have to apply a pressure in the opposite direction to bring the mercury in the capillary to its level in the vessel. Clearly, the physics controlling capillary rise and capillary depression must be the same. To accommodate the difference between the two in the framework of the same equations, we note the difference in the form of the meniscus in Figure 9.2a,b. For water, the meniscus is concave, while for mercury it is convex. If we agree to define

the radius as being *negative* for the convex interphase, and choose the level of the liquid in the outer vessel as zero (leading to negative values of h for capillary depression), we find that the Young–Laplace equation is applicable to wetting as well as non-wetting liquids, and Eq. (9.18) correctly describes both phenomena.

The reader may object to the use of an artificial concept such as a negative radius, but this is not an uncommon practice in science. For instance, the capacitive impedance of the metal/solution interphase is described as the "imaginary" part of the impedance, although it is a very real impedance indeed!

We have obviously taken some shortcuts to keep the foregoing discussion short and, hopefully, clear. For example, we ignored the angle of contact between the liquid and the solid at the edge of the meniscus. This is tantamount to considering water to be an *ideally wetting* liquid (with a contact angle of zero), and mercury to be an *ideally non-wetting* liquid (having a contact angle of 180°). One may also question the exact definition of the height h, in Figure 9.2a. Is it measured to the bottom of the meniscus, to the point at which the meniscus contacts the glass, or somewhere in between? Is the capillary ideally cylindrical, leading to a hemispherical meniscus, or should we define two radii of curvature? These are important points when one conducts research in this field, but they are not relevant to the basic physical understanding of the phenomena observed. Although Figure 9.2 is not drawn to scale, the capillary rise is intentionally drawn larger than the capillary depression, to show the effect of the high density of Hg (see Eq. (9.16)).

In the classical electrocapillary electrometer, the configuration is inverted. Mercury is placed in a glass tube that ends with a fine capillary, as shown in Figure 9.3. Since we need pressure to force mercury into a fine capillary, there will be a certain height of mercury column supported by the capillary in this configuration. This is the exact equivalent of the capillary depression shown in Figure 9.2b, and the height of the column is also given by Eq. (9.18). In this equation, we note that the height depends on the surface tension, which depends on potential. Hence, the height of the mercury column above the capillary is a function of potential.

Measurement of the electrocapillary curve involves changing the potential stepwise and determining the pressure required to return the mercury meniscus to the same location in the fine capillary. A plot of this pressure as a function of potential is nothing but the electrocapillary curve, multiplied by a constant, according to Eq. (9.18). Although this equation allows an absolute determination of the surface tension, the best way to determine it is by calibration with a known system. This requires one accurate determination of σ by an independent method. Very careful experiments were performed by Gouy at the beginning of the twentieth century. The highest value he measured at the electrocapillary maximum in a solution of NaF is $0.426\,\mathrm{N\,m^{-1}}$. This value is used even now as the *primary standard* for electrocapillary measurements. There are several ways to measure the electrocapillary curve. One of them is shown schematically in Figure 9.3.

The potential of the mercury electrode is controlled against a suitable reference electrode. It is not essential to use a three-electrode system in this case, since the mercury/solution interphase is studied in the range of potentials where it

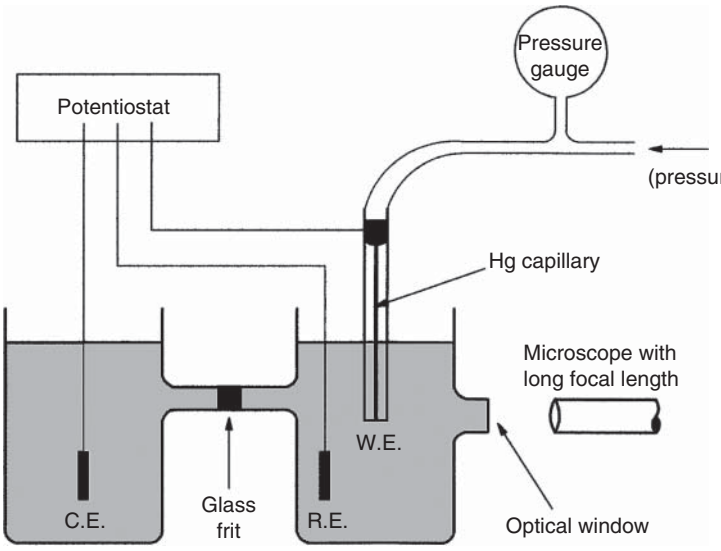


Figure 9.3 The electrocapillary electrometer. For each potential the pressure of nitrogen is adjusted to bring the mercury inside the capillary back to the same fixed position.

behaves essentially as an ideally polarizable interphase. A reference electrode is commonly used, nevertheless, to ensure stability and reproducibility of the measured potential. The position of the mercury in the capillary is observed either by a long-focal-length microscope, or by a video camera connected to a high-resolution monitor. We shall not dwell on technical details, except for noting that building and operating an electrocapillary electrometer at the desired level of accuracy is a delicate matter, requiring both skill and experience.

Let us make a small detour here to discuss a minor point, which appears to be purely technical, but may help us to better understand the physics behind electrocapillary measurements. The question we want to address is: should one make an effort to use a perfectly cylindrical capillary, or is a slightly tapered capillary satisfactory? The answer cannot be found in Eq. (9.18), which relates the height of the mercury column, or the pressure difference, to the radius. Consider, however, the situation in a perfectly cylindrical tube. If we move the meniscus to a different position, it will stay there, since the radius has not been changed. Thus, the system is at equilibrium with the mercury meniscus *anywhere* in the cylindrical tube. What we have is a *neutral* equilibrium. Its mechanical equivalent is a perfect sphere on a perfectly flat and horizontal surface, as shown in Figure 9.4a.

Now, consider a tapered tube, wider at the top. This represents a stable equilibrium. For a given pressure applied, the mercury is stable in only one position in the capillary. If the pressure is increased momentarily (or the surface tension is decreased, say, by a fluctuation in the applied potential), the mercury meniscus will move to a lower point, where the radius is a little smaller, to establish a new equilibrium, in accordance with Eqs. (9.16)–(9.18). The mechanical equivalent of this configuration is a sphere at the bottom of a concave surface, shown in Figure 9.4b. Stability is attained by negative feedback. If the system is perturbed,

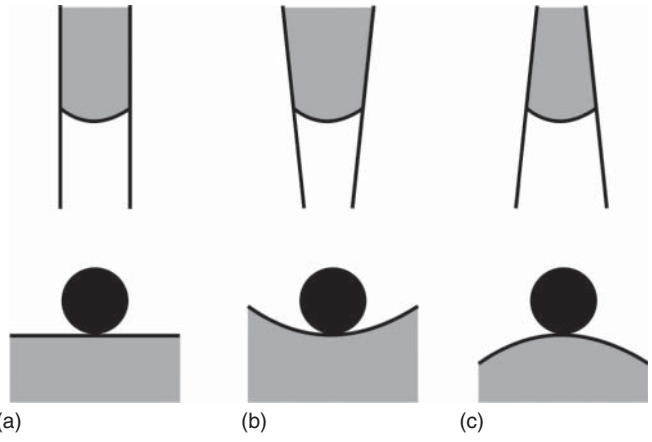


Figure 9.4 Neutral, stable, and unstable equilibria (a–c, respectively) in mechanics and in an electrocapillary electrometer.

a force is created, acting in the direction opposite to the direction of the perturbing force, bringing it to a new equilibrium position.

Consider now the same tapered capillary, but placed in the inverted position, namely, wider at the bottom. This represents an *unstable equilibrium*. The mechanical equivalent is a sphere on the top of a convex surface, as shown in Figure 9.4c. It is theoretically possible to find an equilibrium position for the sphere, but it will be impossible to maintain it in practice. The slightest disturbance will cause it to roll down. Unstable equilibrium is the result of positive feedback. Removing the system even very slightly from its state of equilibrium creates a force in the direction of removing it even further away from equilibrium.

What is the effect of a small increase in pressure on the mercury in such an inverted capillary? The meniscus will move downward slightly, but in its new position the radius is larger, so that the same pressure will tend to move it farther down, where the radius is even larger, and so on. Thus, while the general equations allow the existence of an equilibrium position in this situation, it will not be possible in practice to maintain a stable position of the mercury meniscus in the capillary.

Fortunately, the normal manufacturing process of glass tubing produces slightly tapered capillaries, and the problem does not arise, unless one connects the capillary in the electrocapillary electrometer in the wrong direction.

The information one can derive from measurements of the surface tension as a function of potential is specified in Eqs. (9.10)–(9.15). It includes the dependence of charge and the double-layer capacitance on the potential, and the relative surface excess of all the species in solution (except the solvent) as a function of the potential and the composition of the solution. It should be noted, however, that all these quantities must be obtained by numerical differentiation of the experimental results. This requires very high accuracy in measurement, since differentiation inherently amplifies experimental errors (while integration tends to smooth them

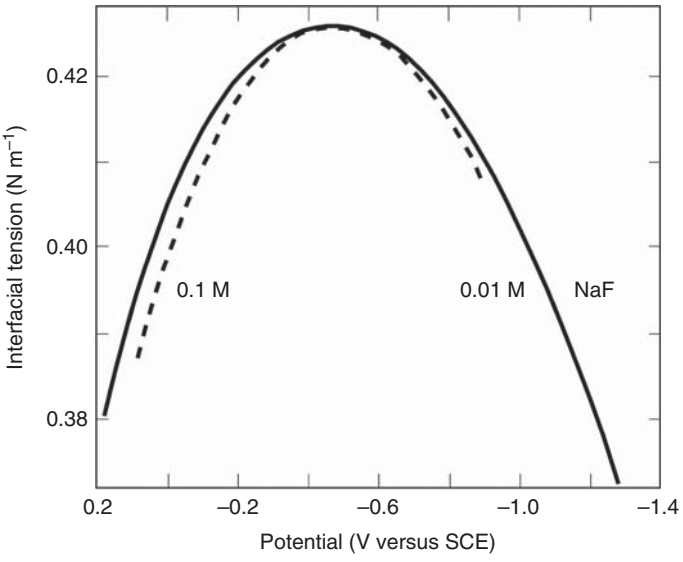


Figure 9.5 Electrocapillary curves for two concentrations of NaF. Source: Data from D.C. Grahame and B.A. Soderberg, *Tech. Rep. Nº* 14, ONR, 1954.

out). The range of values of σ in most cases is 0.250–0.426 N m^{-1}. The best measurements recorded claim an accuracy of $\pm 10^{-4}$ N m^{-1}, which amounts to about $\pm$ (0.04–0.02)%.

9.2.2 Some Experimental Results

9.2.2.1 The Adsorption of Ions

Figure 9.5 shows electrocapillary curves obtained in NaF solutions at different concentrations. The most remarkable feature of these curves is that both E_{pzc} and σ_{max} are almost independent of the concentration of the electrolyte. Also, there is hardly any adsorption of the anion at negative potentials, or of the cation at positive potentials. This observation led Grahame and others to conclude that neither Na^+ nor F^- are specifically adsorbed in the mercury/electrolyte interphase.

A distinctly different behavior is shown in Figure 9.6 for KBr. The potential of zero charge is shifted in the negative direction as the concentration of the electrolyte increases. The anion is strongly adsorbed, as can be seen from the depression of the value of γ with increasing concentration, at fixed potential, which is observed slightly even on part of the negative branch of the electrocapillary curve.

A case of specific adsorption of the cation is shown in Figure 9.7, for solutions of $TlNO_3$ in a supporting electrolyte consisting of 1.0 M KNO_3 and 0.01 M HNO_3. The potential of zero charge shifts in the positive direction, showing that specific adsorption of the cation is predominant. The values of the surface tension are substantially lowered at all potentials, indicating high values of the surface excess of thallium.

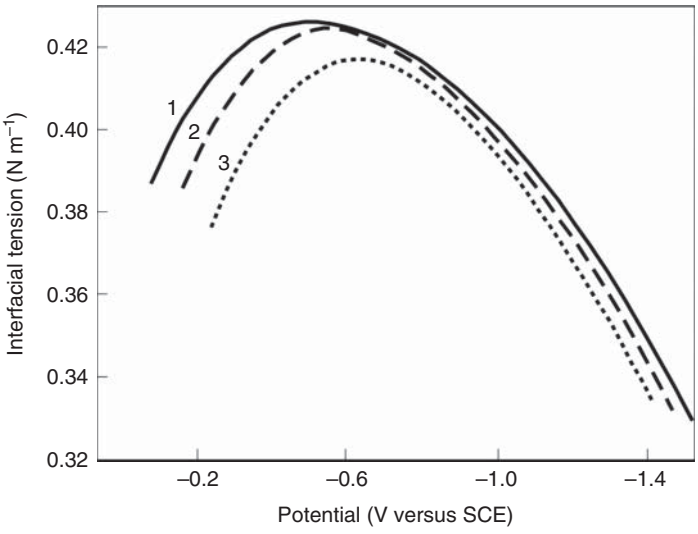


Figure 9.6 Electrocapillary curves for different concentrations of KBr. Lines 1, 2, and 3 for 0.01, 0.1, and 1 M KBr, respectively. Source: Based on data from J. Lawrence, R. Parsons, and R. Payne *J. Electroanal. Chem*. **16** (1968) 193.

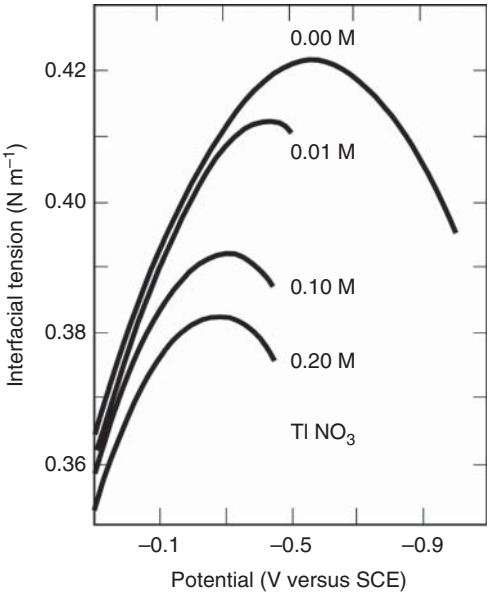


Figure 9.7 Electrocapillary curves for solutions of $TlNO_3$ in 1.0 M KNO_3 with 0.01 M HNO_3. Source: Reprinted with permission from A.N. Frumkin, *Trans. Symp. Electrode Processes*, 1961, pp. 1–15.

9.2.2.2 Adsorption of Neutral Molecules

The adsorption of neutral molecules is discussed in Chapter 12. Here we show electrocapillary curves obtained in the presence of *n*-butanol and *n*-C_4H_9O, and compare them to an electrocapillary curve taken in a solution of NaF. Looking at Figure 9.8 we note that adsorption occurs mainly in the vicinity of E_{pzc}, and the

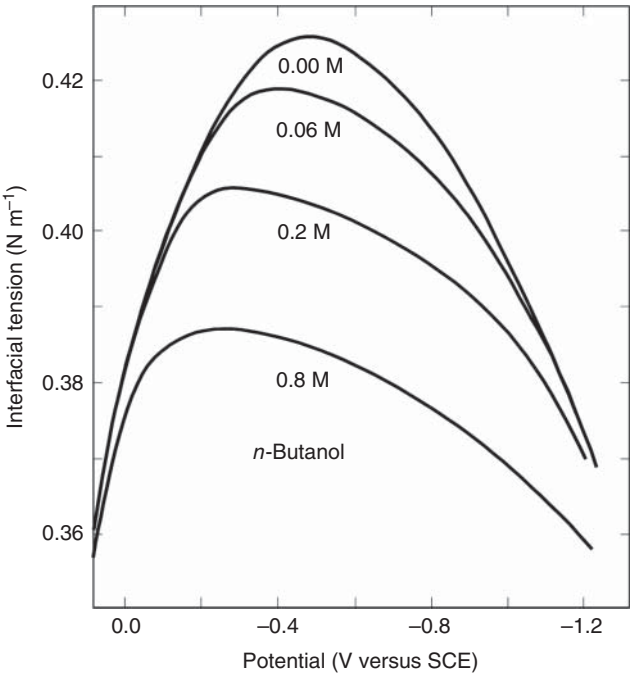


Figure 9.8 Electrocapillary curves in the presence of different concentrations of *n*-butanol in 0.1 M HCl. Source: Based on data by J.O'M. Bockris, E. Gileadi, and K. Müller, *Electrochim. Acta* **12** (1967) 1301.

organic substance seems to be "pushed out" of the interphase at both negative and positive values of the potential.

This behavior is not limited to the compound shown in Figure 9.8, nor can it be related to the structure of the adsorbed molecule or its dipole moment. Very similar behavior has been observed, for example, for benzene and for phenol, although the former has no permanent dipole moment, whereas the dipole moment of phenol is substantial. Results similar to those shown in Figure 9.8 were also obtained for *n*-butyl-cyanide (n-C_4H_9CN), although its dipole moment is quite different from that of *n*-butanol. It is logical, then, to conclude that the observed behavior is primarily a property of the solvent. Thus, the adsorption of an organic molecule requires the removal of a certain number of water molecules (depending on the size of the molecule being adsorbed), and it is the effect of the potential on the energy required to remove these water molecules that determines, indirectly, the potential dependence of adsorption of neutral species. The adsorption of a neutral molecule and the simultaneous removal of an appropriate number of water molecules from the interphase is called *electrosorption*. This is discussed in detail in Chapter 12.

Another interesting feature of the curves plotted in Figure 9.8 is the relatively sudden convergence of the curves in the presence and in the absence of butanol at high positive and high negative potentials, corresponding to a sudden transition from strong adsorption to practically no adsorption. This leads to low

values of the double-layer capacitance in the region of adsorption, and to sharp *adsorption–desorption peaks* in the plot of C_{dl} versus E, which are shown in Figure 8.2.

Further Reading

1 Soustelle, M. (2016). *Thermodynamics of Surfaces and Capillary Systems*, Chemical Engineering Series, vol. 7. London, UK: ISTE and Wiley.
2 Chakraborty, S. (2014). Electrocapillary. In: *Encyclopedia of Microfluidics and Nanofluidics* (ed. D. Li), 460–469. New York, USA: Springer Science+Business Media.
3 Newman, J. and Thomas-Alyea, K.E. (2004). *Electrochemical Systems*, 3rd edn. New Jersey, USA: Wiley.
4 Bard, A.J. and Faulkner, L.R. (2001). *Electrochemical Methods: Fundamentals and Applications*, 2nd edn. New York, USA: Wiley.
5 Valincius, G. (1999). Electrocapillary equations of solid electrodes. *J. Electroanal. Chem.* 478: 40–49.
6 Gileadi, E. (1993). *Electrode Kinetics – for Chemists, Chemical Engineers and Materials Scientists*. Wiley-VCH.
7 Gileadi, E., Kirowa-Eisner, E., and Penciner, J. (1975). *Interfacial Electrochemistry – An Experimental Approach*. Addison-Wesley.
8 Lawrence, J., Parsons, R., and Payne, R. (1968). Adsorption of halides at the mercury-water interface. *J. Electroanal. Chem.* 16: 193–206.
9 Muller, K. (1965). On the structure of charged interfaces. Ph.D. Dissertation. University of Pennsylvania.
10 Frumkin, A.N. (1961). The adsorption of ions at the metal–solution interface and its influence on electrode kinetics. In: *Transactions of the Symposium on Electrode Processes* (ed. E. Yeager), 1–15. New York, NY, USA: Wiley.
11 Grahame, D.C. and Soderberg, B.A. (1954). Thermodynamic properties of the double layer. *Technical Report No. 14*. Office of Naval Research, Amherst College.

10

Intermediates in Electrode Reactions

10.1 Adsorption Isotherms for Intermediates Formed by Charge Transfer

10.1.1 General

An adsorption isotherm can be written in general form as

$$f(\theta) = g(c_b) \tag{10.1}$$

It represents the equilibrium between a species in solution and the same species adsorbed on the surface. The fractional surface coverage is defined as

$$\theta \equiv \Gamma/\Gamma_{max} \tag{10.2}$$

This definition implies that $0 \leq \theta \leq 1$, namely that adsorption is limited to a monolayer, which is valid if the bonding of the first layer of the adsorbate to the substrate is much stronger than that of subsequent layers to each other. Multilayer adsorption can occur under suitable conditions, but will not be discussed here.

10.1.2 The Langmuir Isotherm and its Limitations

The *Langmuir isotherm* is written in chemistry in the form

$$\frac{\theta}{1-\theta} = Kc_b \tag{10.3}$$

where K is the equilibrium constant for adsorption. The standard Gibbs free energy of adsorption is related to the equilibrium constant as in every chemical equilibrium, namely

$$2.3RT \log K = -\Delta G^0_{ads} \tag{10.4}$$

The Langmuir isotherm is only applicable if the following conditions are satisfied:

(i) The Gibbs free energy of adsorption is independent of coverage. This is a severe limitation! For this condition to apply, it must be assumed that the surface is completely homogeneous, and there are no lateral interactions among the adsorbed species. Solid surfaces are rarely homogeneous on the atomic scale. There are highly active sites and others having low activity. As

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

the fractional surface coverage increases, the most active sites are occupied first, and the least active ones are occupied last. When this is the case, the absolute value of the standard Gibbs free energy of adsorption decreases with increasing coverage.

The purists among us may object to the last statement. A standard Gibbs free energy is defined for some standard state. From a formal point of view, it cannot be said to depend on any variable. Yet, if we think of a heterogeneous surface as consisting of a very large number of minute homogeneous patches, we would have a different standard Gibbs free energy for each little patch, and the experimentally observed standard Gibbs free energy of adsorption would, in effect, be a function of coverage. To overcome this problem, it is customary to discuss the variation of the apparent standard Gibbs free energy of adsorption with coverage. This distinction does not change the ensuing arguments, but it helps to soothe our scientific conscience. For brevity, we shall drop the word "apparent" in the following discussion.

What about lateral interactions? These depend on the nature of the adsorbed species and on their average distance apart. Ion–ion interactions are long-range because ions are, by definition, charged. The interaction with their neighbors is electrostatic, and the energy of interaction decays with the first power of the distance. Dipole–dipole interactions decay with r^{-3}, while chemical interactions decay with r^{-6}, and are expected to be felt only at higher values of the coverage. But, what will be considered a high coverage? In solution chemistry, a concentration of $c_b \leq 10\,\text{mM}$ is usually regarded to be a dilute solution, where interactions among solute molecules is small, while for $c_b \geq 1\,\text{M}$ the solution would be considered to be concentrated, so that such interactions can usually not be ignored. In electrode kinetics, values of $\theta \leq 0.1$ and $\theta \geq 0.9$ are considered to represent low and high coverage, respectively. This point will be discussed further in the next section.

(ii) The surface is assumed to have well-defined sites, with each molecule of the adsorbate occupying a single site, and each site able to accommodate only a single adsorbate molecule. For example, the adsorption of Ag on Au satisfies this condition, since the crystal radii of the two elements are almost equal (Ag: 0.1445 nm, Au: 0.1442 nm). On the other hand, the adsorption of Pb on Au cannot be fitted to the Langmuir isotherm, since the atomic radius of Pb is 0.175 nm, so that it is impossible to adsorb an atom of Pb on each site. It follows also that adsorption of large molecules, each occupying several sites on the surface, cannot be treated by applying the Langmuir isotherm.

(iii) Equilibrium between the species in the bulk and on the surface is assumed.

(iv) The isotherm is applicable only to monolayer adsorption. The last point may seem obvious when an isotherm is discussed in the context of adsorption of intermediates in electrode kinetics, but it should be emphasized here that this does not apply to the discussion of surface excess and the Gibbs adsorption isotherm, which was treated in Chapter 9.

10.1.3 Application of the Langmuir Isotherm for Charge-Transfer Processes

Consider the application of the Langmuir isotherm to electrochemistry. For a simple charge-transfer process, such as the formation of an adsorbed hydrogen atom, we write

$$H_3O^+ + e_M^- + Pt^0 \rightleftharpoons Pt\text{-}H + H_2O \tag{10.5}$$

The corresponding isotherm is

$$\left[\frac{\theta}{1-\theta}\right] = K c_b \exp\left(\frac{F}{RT}E\right) \tag{10.6}$$

in which c_b is the bulk concentration of the H_3O^+ ions. Plots of θ versus c_b and versus $\log c_b$ (or versus E) are shown in Figure 10.1, for three values of the equilibrium constant K. Changing the equilibrium constant causes a change in the shape of the plot of θ versus c_b, whereas it causes only a parallel shift in the plot of θ versus $\log c_b$. Plots of θ versus E are similar to those of θ versus $\log c_b$. Changing the potential by $2.3RT/nF$, where n is the number of electrons transferred per atom adsorbed on the surface, has the same effect on θ as changing the concentration by a factor of 10. This is not surprising, in view of the dependence of the electrochemical potential on concentration and potential, given by

$$\overline{\mu}_i = \mu_i^0 + 2.3RT \log c_i + n_i \phi F \tag{10.7}$$

The Langmuir isotherm can be used in electrode kinetics under the special conditions of θ approaching either zero or unity. This relieves us of the most difficult assumption, namely that the standard Gibbs free energy of adsorption is independent of coverage. Thus, even if ΔG_{ads}^0 changes significantly with θ, this change can be considered to be negligible, when the coverage is small ($\theta \leq 0.1$), or when it is large ($\theta \geq 0.9$). It is not surprising then, that the Langmuir isotherm is often applicable in electrode kinetics, in spite of its many limitations. There are, however, cases in which it is not applicable, as discussed in the next section.

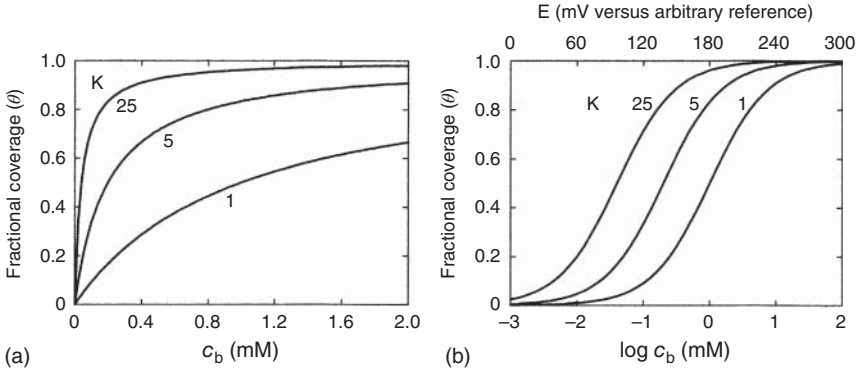


Figure 10.1 The Langmuir isotherm. The fractional surface coverage (a) versus bulk concentration and (b) versus E or $\log c_b$ for three values of the equilibrium constant $K = 1$, 5, and 25.

10.1.4 The Frumkin Adsorption Isotherms

The assumption that the standard Gibbs free energy of adsorption is independent of coverage may be viewed as a first-order approximation. A better approximation will then be to assume a linear dependence of the form

$$\Delta G_\theta^0 = \Delta G_0^0 + r\theta \tag{10.8}$$

where ΔG_0^0 and ΔG_θ^0 are the standard Gibbs free energies of adsorption for $\theta = 0$ and for a chosen value of θ, respectively. The parameter r is the rate of change of the standard Gibbs free energy of adsorption with coverage. Considering that for the adsorption process both ΔG_θ^0 and ΔG_0^0 are negative, it may be better to write Eq. (10.8) as

$$|\Delta G_\theta^0| = |\Delta G_0^0| - r\theta \tag{10.9}$$

showing that the *absolute value* of the standard Gibbs free energy of adsorption decreases with increasing coverage.

Keeping in mind the relationship between the standard Gibbs free energy and the equilibrium constant, it is very easy to modify the Langmuir isotherm, taking into account the variation of ΔG_{ads}^0 with θ. To do this, we replace ΔG_{ads}^0 in Eq. (10.4) by ΔG_θ^0 and combine with Eq. (10.8), to obtain

$$2.3RT \log\ K = -\Delta G_\theta^0 = -(\Delta G_0^0 + r\theta) \tag{10.10}$$

which can also be written as

$$K = \exp\left(-\frac{\Delta G_0^0 + r\theta}{RT}\right) = K_0 \exp\left(-\frac{r}{RT}\theta\right) = K_0 \exp\ (-f\theta) \tag{10.11}$$

where $K_0 = K \exp(-\Delta G_0^0/RT)$ and the Frumkin parameter, f, is defined as

$$f \equiv \frac{r}{RT} \tag{10.12}$$

Substituting this value of K into Eq. (10.6) yields

$$\frac{\theta}{1-\theta} \exp\left(\frac{r}{RT}\theta\right) = K_0\, c_\text{b} \exp\left(\frac{F}{RT}E\right) \tag{10.13}$$

This equation is known as the *Frumkin isotherm*. It should be clear that the Langmuir isotherm is a special case of the Frumkin isotherm, derived from it, if $r = 0$. It can also be seen that, for reasonable values of the parameter r (in the range of 20–60 kJ mol^{-1}), the exponential term on the left-hand side of this equation approaches unity for very small values of θ, and becomes constant when θ is close to unity. Thus, at extreme values of θ, the Frumkin and the Langmuir isotherms lead to the same dependence of coverage on potential, and hence to the same rate equations in electrode kinetics.

The dependence of θ on potential is shown in Figure 10.2 for four values of the parameter f. What can be learned from Figure 10.2 and Eq. (10.13)? First, we note that θ increases more slowly with potential as the value of the parameter f is increased. For example, it takes only 0.11 V to change the coverage from 0.1 to 0.9, if the Langmuir isotherm applies ($f = 0$). The same increase in θ occurs over about 0.52 V for the Frumkin isotherm, for $f = 20$.

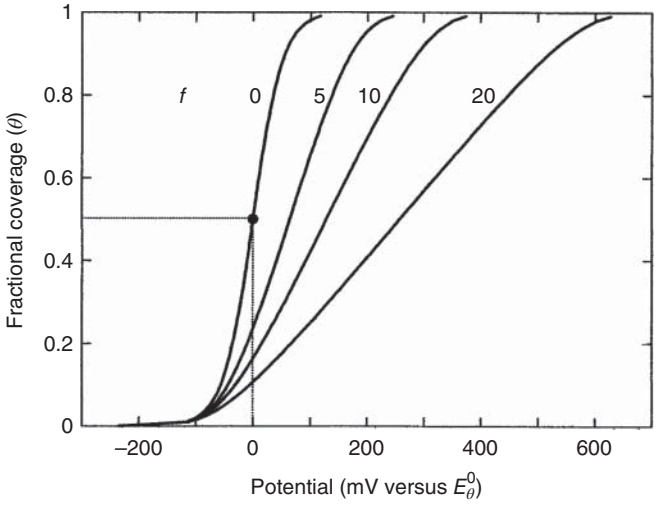


Figure 10.2 The dependence of θ on E for different values of the Frumkin parameter ($f \equiv r/RT$). The dotted lines on the curve for $f = 0$ meet at $\theta = 0.5$ and $E = E^0_\theta$.

Secondly, at intermediate values of the coverage, the pre-exponential term $\theta/(1-\theta)$ varies little with θ compared to the variation of the exponential term. Thus, taking $\theta/(1-\theta) \approx 1$, the Frumkin isotherm can be written approximately as

$$\exp\left(\frac{r}{RT}\theta\right) = K_0 c_\mathrm{b} \exp\left(\frac{F}{RT}E\right) \tag{10.14}$$

or, in logarithmic form

$$\theta = (2.3RT/r)\log\ (K_0 c_\mathrm{b}) + \frac{F}{r}E \tag{10.15}$$

Equation (10.15) shows a linear dependence of θ on E for intermediate values of the fractional coverage, as seen in Figure 10.2. In the same intermediate region of coverage, Eq. (10.15) also shows a logarithmic dependence of the coverage on the bulk concentration. A similar *logarithmic isotherm* was developed independently by Temkin. Consequently, Eq. (10.15) is referred to in the field of electrode kinetics as the *Temkin isotherm*, characterized by a linear dependence of coverage on potential at intermediate value of the partial coverage θ.

10.2 The Adsorption Pseudocapacitance C_ϕ

10.2.1 Formal Definition of C_ϕ and its Physical Understanding

The adsorption isotherms discussed in Section 10.1 above describe the potential dependence of the fractional surface coverage, θ. For intermediates formed

in a charge-transfer process (see Eq. (10.5)), the fractional surface coverage is associated with a faradaic charge q_F that is given by

$$q_F = nF\Gamma \tag{10.16}$$

where Γ is the surface concentration, corresponding to any given value of θ. If we denote the charge required to form a complete monolayer of a species by q_1, we have the simple relationship

$$q_F = q_1\theta \tag{10.17}$$

Thus, the adsorption isotherm also yields the dependence of the faradaic charge consumed in forming the adsorbed intermediate on potential. This allows us to define a new type of capacitance, which is called the *adsorption pseudocapacitance,* C_ϕ as

$$C_\phi \equiv (\partial q_F/\partial E)_{\mu_i} = q_1(\partial\theta/\partial E)_{\mu_i} \tag{10.18}$$

This rather important concept in electrochemistry warrants some clarification. Evidently, we are not dealing here with a pure capacitor, such as C_{dl}, because charge transfer is involved. Moreover, the very existence of the adsorption pseudocapacitance is linked to charge transfer. In contrast, the double-layer capacitance behaves as a pure capacitor: when charge is *brought* to one side of the capacitor, an equal but opposite charge is *induced* on the other side. An excess of electrons on the surface of the metal causes a rearrangement of the distribution of ions on the solution side of the interface, yielding an excess of positively charged ions, and vice versa. There is no transfer of charge across the interface.

Although not a pure capacitor, the adsorption pseudocapacitance exhibits some of the properties typical of capacitors. Whatever the type of isotherm applicable to the system, there is a singular relationship between the charge and the potential. Setting the potential determines the charge, and vice versa. When the potential is changed, a transient (faradaic) current is observed. The current decays to zero when the charge passed is enough to bring the fractional coverage from its initial value to the value corresponding to the new potential. At a fixed potential, the steady-state current is zero. This is exactly the way a pure capacitor should behave. It allows the passage of transient currents, but presents an infinite resistance to direct current.

The capacitive nature of the adsorption pseudocapacitance can be further illustrated by considering its response to an alternating voltage (AC) perturbation. Let us assume that a low-amplitude sinusoidal voltage signal is applied to a system at equilibrium. The sinusoidal waveform can be expressed by the equation

$$E = E_{DC} + (\Delta E)\sin(\omega t) \tag{10.19}$$

where E_{DC} is the potential attributed to the direct current maintained, ΔE is the amplitude of the sine wave, and ω is the angular velocity, related to the frequency as $\omega = 2\pi\nu$, where ν is the frequency. We assume here that the frequency is low enough, so that the coverage at any moment is equal to its equilibrium value, corresponding to the momentary value of the potential, as shown in Figure 10.3.

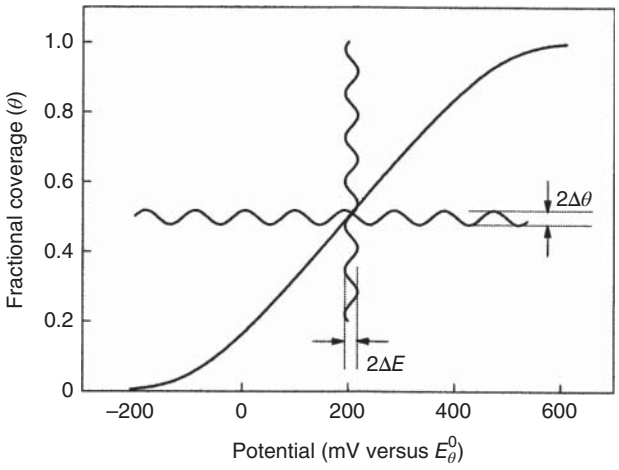


Figure 10.3 The variation of the coverage around $\theta = 0.5$, as a result of a small AC perturbation of the potential. The Frumkin isotherm with an interaction parameter $f = 20$ is used here, for illustration.

The rate of change of coverage with time is proportional to the rate of change of potential with time:

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} \propto \frac{\mathrm{d}E}{\mathrm{d}t} = (\omega\,\Delta E)\cos(\omega t) \tag{10.20}$$

The faradaic current is given by

$$j_\mathrm{F} = \frac{\mathrm{d}q_F}{\mathrm{d}t} = q_1\frac{\mathrm{d}\theta}{\mathrm{d}t} \tag{10.21}$$

which can also be written as

$$j_\mathrm{F} = (\Delta j_\mathrm{F})\cos(\omega t) = (\Delta j_\mathrm{F})\sin(\omega t - \pi/2) \tag{10.22}$$

where Δj_F is the amplitude of the AC current. The phase retardation of $-\pi/2$ between potential and current is the expected response of a pure capacitor to a sinusoidal voltage perturbation.

10.2.2 The Equivalent-Circuit Representation

What makes C_ϕ a *pseudocapacitance*, rather than a regular capacitance? Obviously, it is the fact that it is intimately related to, and indeed dependent on, charge transfer across the interface. The equivalent circuit that represents the adsorption pseudocapacitance itself is a resistor and a capacitor connected in series, unlike the double-layer capacitance, C_dl, which is represented as a pure capacitance, as shown in Figure 10.4.

Note that the resistance R_ϕ is an integral part of the physical phenomenon that gives rise to the formation of the adsorption pseudocapacitance. It is a faradaic resistance, since C_ϕ is due to a charge-transfer process. The association of this charge-transfer process with the formation of an adsorbed intermediate, which can proceed only until the appropriate coverage has been reached, is manifested

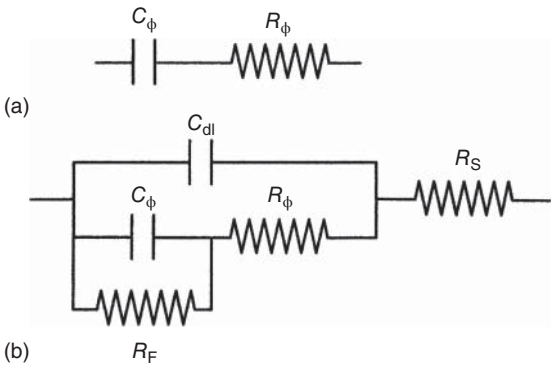


Figure 10.4 The equivalent circuits for (a) the adsorption pseudocapacitance C_ϕ and the corresponding resistance R_ϕ. (b) An interface containing an adsorption pseudocapacitance.

by placing the resistor in series with the capacitor. It should also be borne in mind that both C_ϕ and R_ϕ may depend on the potential, as in other equivalent circuits representing the electrochemical interface.

How is the adsorption pseudocapacitance affected by frequency? Returning to Figure 10.3, we recall that it was assumed that the change of θ can track the change of potential during the measurement. In other words, it was assumed that at every point in time, the value of θ is equal to its equilibrium value at that potential. But for this to happen, the charge-transfer reaction must occur fast enough, and the rate of this reaction is proportional to $1/R_\phi$. As the frequency is increased, the changes in partial coverage can no longer keep up with the changes in potential. Eventually, at a sufficiently high frequency, the adsorbed intermediate is effectively "frozen in" – the potential changes back and forth so fast that the coverage does not have a chance to follow. Since C_ϕ is proportional to $\mathrm{d}\theta/\mathrm{d}E$, the adsorption pseudocapacitance tends to zero when the time constant $\tau_\phi = C_\phi \times R_\phi$ of the circuit shown in Figure 10.4a is much longer than the period of the AC signal.

10.2.3 Calculation of C_ϕ as a function of θ and E

The adsorption pseudocapacitance can be readily calculated from the appropriate isotherm, with the use of its definition given in Eq. (10.18). This is shown next for the Langmuir and Frumkin isotherms.

The *Langmuir isotherm* describing the formation of an adsorbed intermediate by charge transfer (see Eq. (10.6)) can be written in a different form as

$$\theta = \frac{K c_\mathrm{b} \exp(EF/RT)}{1 + K c_\mathrm{b} \exp(EF/RT)} \tag{10.23}$$

from which we obtain

$$\frac{\mathrm{d}\theta}{\mathrm{d}E} = (F/RT)\left[\frac{K c_\mathrm{b} \exp(EF/RT)}{[1 + K c_\mathrm{b} \exp(EF/RT)]^2}\right] \tag{10.24}$$

Combining with Eq. (10.17) we get

$$C_\mathrm{L} = q_1 \frac{\mathrm{d}\theta}{\mathrm{d}E} = \frac{q_1 F}{RT}\left[\frac{K c_\mathrm{b} \exp(EF/RT)}{[1 + K c_\mathrm{b} \exp(EF/RT)]^2}\right] \tag{10.25}$$

where C_L is the adsorption pseudocapacitance derived from the Langmuir isotherm. Taking the second derivative with respect to potential, it can be shown that the maximum value of the adsorption pseudocapacitance is reached when $K c_b \exp(FE/RT) = 1$, corresponding to $\theta = 0.5$. Thus, we can write

$$C_L(\max) = \frac{q_1 F}{4RT} \quad \text{and} \quad E_{\max} = -\frac{2.3RT}{F} \log(Kc_b) \tag{10.26}$$

where $E_{\max}$ is the potential where the pseudocapacitance reaches its maximum value.

The dependence of C_L on coverage can be obtained by rewriting the Langmuir isotherm in the form

$$E = \frac{2.3RT}{F} \log\left[\frac{\theta}{1-\theta}\right] - \frac{2.3RT}{F} \log(Kc_b) \tag{10.27}$$

Differentiating with respect to charge one obtains

$$\frac{1}{C_L} = \left(\frac{1}{q_1}\right)\left(\frac{dE}{d\theta}\right) = \left(\frac{RT}{q_1 F}\right)\left[\frac{1}{\theta(1-\theta)}\right] \tag{10.28}$$

which can be rearranged to

$$C_L = \frac{q_1 F}{RT}\theta(1-\theta) \tag{10.29}$$

Clearly, C_L has its maximum value, given by Eq. (10.26), at $\theta = 0.5$. Setting the concentration in Eq. (10.26) equal to unity, we note that the potential at which C_L is at maximum can be regarded as the standard potential E_θ^0 for the adsorption process. It is given by

$$E_\theta^0 = -\frac{2.3RT}{F} \log K \tag{10.30}$$

The physical meaning of this choice is that the standard state of the system is chosen to be $\theta = 0.5$ and $c_b = 1.0$ M.

It is interesting to evaluate the numerical value of $C_{L,\max}$. Taking a value of $q_1 = 0.23$ mC cm^{-2} for a monolayer of single-charged species, we find that

$$C_{L,\max} = \frac{0.23 \times 10^{-3} \times 96.5 \times 10^3}{4 \times 8.31 \times 298} = 2.24 \text{ mF cm}^{-2} \tag{10.31}$$

This is more than 2 orders of magnitude higher than the typical values observed for the double-layer capacitance. We can also evaluate the range of potential over which $C_L/C_{dl} \geq 1$. Substituting $C_L = 16$ μF cm^{-2} in Eq. (10.31), we find that the two values of θ that satisfy this relationship are 0.00187 and 0.99813. In other words, $C_L/C_{dl} \geq 1$ for values of θ between about 0.2% and 99.8%. The corresponding range of potential (found by introducing these values of θ into the Langmuir isotherm) is 0.16 V.

The important conclusion to be drawn from this numerical calculation is that great care must be exercised in interpreting double-layer capacitance measurements in systems in which an adsorbed intermediate can be formed. Even a minute fractional surface coverage can give rise to an adsorption pseudocapacitance comparable to or larger than the double-layer capacitance.

We proceed here to derive the expressions for the adsorption pseudocapacitance when the *Frumkin isotherm*, which we shall denote as C_F, applies. Adding the interaction parameter to Eq. (10.13) we obtain the Frumkin isotherm

$$E = \frac{RT}{F}\ln\left[\frac{\theta}{1-\theta}\right] + \frac{r}{F}\theta - \frac{RT}{F}\ln(K_0 c_b) \tag{10.32}$$

We write the parameter r in dimensionless form as $f \equiv r/RT$. Differentiating Eq. (10.32) yields

$$\frac{dE}{d\theta} = \frac{RT}{F}\left[\frac{1}{\theta(1-\theta)}\right] + \frac{RT}{F}f = \frac{RT}{F}\left[\frac{1+f\theta(1-\theta)}{\theta(1-\theta)}\right] \tag{10.33}$$

Thus, the adsorption pseudocapacitance is derived for the Frumkin isotherm as

$$C_F = \frac{q_1 F}{RT}\left[\frac{\theta(1-\theta)}{1+f\theta(1-\theta)}\right] = C_L\left[\frac{1}{1+f\theta(1-\theta)}\right] \tag{10.34}$$

This equation yields a maximum value of

$$C_{F,max} = \frac{q_1 F}{4RT}\left[\frac{1}{1+f/4}\right] = C_{L,max}\left[\frac{1}{1+f/4}\right] \tag{10.35}$$

The dependence of C_ϕ on θ is shown in Figure 10.5a. We note that all curves are symmetrical with respect to $\nu = 0.5$, irrespective of the value of the parameter f. The maximum declines rapidly with increasing values of the parameter f. The dependence of the adsorption pseudocapacitance on potential is shown in Figure 10.5b. This function cannot be derived directly, since it is impossible to express θ as an explicit function of E from Eq. (10.15). However, since the dependence of both E and C_F on θ is known, the curves shown in Figure 10.5b can readily be calculated. The peak is sharply decreased with increasing values of the interaction parameter, and in addition, the peak is shifted in Figure 10.5b. It is also noted that the value of the adsorption pseudocapacitance is almost constant, over a fairly large range of potential.

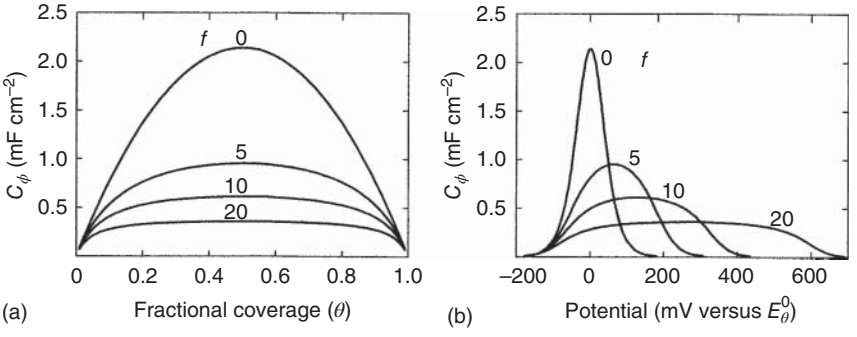


Figure 10.5 The dependence of C_ϕ on (a) coverage and (b) potential for different values of the parameter *f*. $q_1 = 230\ \mu C\ cm^{-2}$.

Further Reading

1 Broussea, T., Bélangerc, D., and Long, J.W. (2015). To be or not to be pseudocapacitive? *J. Electrochem. Soc.* 162: A5185–A5189.
2 Augustyn, V., Simon, P., and Dunn, B. (2014). Pseudocapacitive oxide materials for high-rate electrochemical energy storage. *Energy Environ. Sci.* 7: 1597–1614.
3 Bard, A.J., Inzelt, G., and Scholz, F. (ed.) (2012). *Electrochemical Dictionary*. Berlin Heidelberg: Springer-Verlag.
4 Schmickler, W. and Santos, E. (2010). *Interfacial Electrochemistry*, 2nd edn. Berlin Heidelberg: Springer-Verlag.
5 Volkova-Gugeshashvili, M.I., Volkov, A.G., and Markin, V.S. (2006). Adsorption at liquid interfaces: The generalized Frumkin isotherm and interfacial structure. *Russ. J. Electrochem.* 42: 1073–1078.
6 Newman, J. and Thomas-Alyea, K.E. (2004). *Electrochemical Systems*, 3rd edn. New Jersey, USA: Wiley.
7 Conway, B.E. and Pell, W.G. (2003). Double-layer and pseudocapacitance types of electrochemical capacitors and their applications to the development of hybrid devices. *J. Solid State Electrochem.* 7: 637–644.
8 Bockris, J.O.'.M., Reddy, A.K.N., and Gamboa-Aldeco, M. (2000). *Modern Electrochemistry*, Fundamental of Electrodics, 2nd edn, vol. 2A. New York, USA: Kluwer Academic Publishers.
9 Wieckowski, A. (ed.) (1999). *Interfacial Electrochemistry: Theory: Experiment, and Applications*. New York, USA: Marcel Dekker.
10 Conway, B.E. (1999). *Electrochemical Supercapacitors: Scientific Fundamentals and Technological Applications*. New York, USA: Springer Science+Business Media.
11 Conway, B.E., Birss, V., and Wojtowicz, J. (1997). The role and the utilization of pseudocapacitance for energy storage by supercapacitors. *J. Power Sources* 66: 1–14.
12 Gileadi, E. (1993). *Electrode Kinetics – for Chemists, Chemical Engineers and Materials Scientists*. Wiley-VCH.
13 Bockris, J.O.'.M. and Khan, S.U.M. (1993). *Surface Electrochemistry: A Molecular Level Approach*. New York, USA: Springer Science+Business Media.
14 Gileadi, E. and Conway, B.E. (1964). The behavior of adsorbed intermediates in electrochemical catalysis. In: *Modern Aspects of Electrochemistry*, vol. 3, Chap 5 (ed. J.O.'.M. Bockris and B.E. Conway), 347–455. London, UK: Butterworth.
15 Conway, B.E. and Gileadi, E. (1962). Kinetic theory of pseudo-capacitance and electrode reactions at appreciable surface coverage. *Trans. Faraday Soc.* 58: 2493–2509.
16 Frumkin, A.N. (1961). The adsorption of ions at the metal–solution interface and its influence on electrode kinetics. In: *Transactions of the Symposium on Electrode Processes* (ed. E. Yeager), 1–15. New York, USA: Wiley.

11

Underpotential Deposition and Single-Crystal Electrochemistry

11.1 Underpotential Deposition (UPD)

11.1.1 Definition and Phenomenology

Underpotential deposition (UPD) is a process in which a monolayer of a metal is deposited cathodically on a different metal substrate, in the range of potential that is *positive* with respect to the reversible deposition of the metal in the same solution. Metal deposition is a reduction process that occurs commonly at potentials negative with respect to the reversible potential. Thus, it might appear that underpotential deposition defies the laws of thermodynamics, but careful analysis of the process shows that it does not.

In Figure 11.1 we show the deposition of Pb^{2+}_{soln} ion on a polycrystalline Ag surface. The experiment was conducted by cyclic voltammetry, which will be discussed in detail in Section 14.4. In this method, a triangular potential waveform is applied from a chosen initial potential (which is −0.10 V versus a silver/silver chloride reference electrode for all curves in this figure), and the direction of the sweep is reversed at different potentials (between −0.50 and −0.35 V in this figure). It can be seen that the cathodic limit of potential for curves 1–4 is positive with respect to the reversible potential for deposition of Pb; that is, it is in the UPD region. The cathodic sweep is where formation of a monolayer of Pb (or a fraction of it) takes place, while the anodic sweep is where the process is reversed, and anodic stripping of this layer takes place, according to the reaction

$$Pb^{0}_{M} \rightarrow Pb^{2+}_{soln} + 2e^{-}_{M} \tag{11.1}$$

The variation of potential with time is given by

$$E = E_{in} \pm \nu t \tag{11.2}$$

where ν is the potential sweep rate. Hence, the abscissa can represent either the potential or the time, and the current under the anodic peak could be integrated to yield the corresponding charge transferred, since

$$q = \int_{E_{in}}^{E} j_{\phi} \mathrm{d}t \tag{11.3}$$

where j_{ϕ} is the faradaic current density associated with the formation of a UPD layer. It is noted that lines 1 and 2 in Figure 11.1 form only a partial monolayer,

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

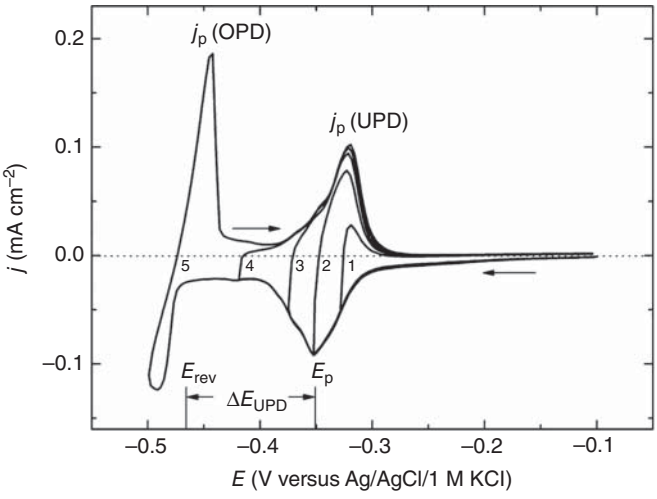


Figure 11.1 Underpotential deposition (UPD) of Pb on polycrystalline Ag, from a solution of 5.0 mM $PbCl_2$ in 0.1 M HCl; $v = 25$ mV s^{-1}. The arrows show the direction of the potential sweep.

while lines 3 and 4 almost overlap, showing that at a potential of −0.375 V (about +0.1 V with respect to the reversible potential), a full monolayer is already formed. Moreover, it is noted that applying a negative overpotential, in the *overpotential* deposition (OPD) region, where bulk deposition occurs, does not change the total charge in the UPD region – it is limited to one monolayer.

How can a metal be deposited at a potential positive with respect to its reversible potential? Deposition of Pb on an electrode made of the same metal can be represented by the equation

$$Pb^{2+}_{soln} + 2e^-_M \rightarrow (Pb - Pb)_M \tag{11.4}$$

while deposition of Pb on another metal (e.g. Ag) should be written as

$$Pb^{2+}_{soln} + 2e^-_M + 2Ag_M \rightarrow (Pb - 2Ag)_M \tag{11.5}$$

The two reactions are similar, but not identical. Eq. (11.4) represents the formation of a bond between a Pb^{2+}_{soln} ion deposited and a Pb_M atom on the surface, while Eq. (11.5) represents the formation of the bond between a Pb^{2+}_{soln} ion deposited on two Ag_M atoms on the surface. Hence, there is no reason a priori to assume that they should have the same reversible potential. Apparently, the Pb–Ag bond is stronger than the Pb–Pb bond, and hence it can be formed at a less negative potential; the laws of thermodynamics have not been violated!

The difference between the potential of the peak current density for formation of a UPD layer and the reversible potential for bulk deposition of the same metal is shown as ΔE_{UPD}. It is given by

$$\Delta E_{UPD} \equiv E_p - E_{rev} \tag{11.6}$$

If this is measured in a solution of standard concentration (1.0 M), and UPD formation is reversible, the potential of the peak current is usually chosen as the standard potential for formation of a UPD layer. It should be emphasized that

ΔE_{UPD} depends on the substrate chosen. For example, its value for the Pb–Ag_2 couple is not the same as that for the Pb–Au_2 couple. This is to be expected, because this potential difference results from the difference between the bonding energy of Pb to Ag and that of Pb to Au. Hence, ΔE_{UPD} should be different for different substrates.

The charge corresponding to a monolayer can readily be determined by employing Eq. (11.3), but it may not be easy to calculate its value a priori. In Chapter 10 we stated that the charge associate with a monolayer of adsorbed intermediates on Pt is about 0.23 mC cm^{-2}, assuming that each adsorbed species transfers a single charge and occupies a single site on the surface.[1] Thus, formation of a full UPD layer of Pb should require a charge of about 0.46 mC cm^{-2}. But, the crystal radii of Pb and Ag are 0.175 and 0.1445 nm, respectively, so that a Pb atom cannot be accommodated on a single site of Ag. Values of q_1 that are between one-third and two-thirds of 0.46 mC cm^{-2} have been reported in the literature for the Ag–Pb system. Moreover, it should be noted that the observed value of q_1 is calculated for the geometrical surface area, ignoring the effect of surface roughness. On the other hand, a monolayer of Cu on Ag, for example, could occupy all sites, because the crystal radius of Cu is only 0.128 nm, smaller than that of Ag.

Figure 11.1 shows the essential features of UPD:

(i) UPD occurs, by definition, at potentials positive with respect to the reversible potential of deposition of the metal in the same solution.

(ii) UPD is a self-limiting process. Deposition is terminated as soon as a complete monolayer has been formed. This statement should be qualified. There are a few cases reported in the literature where more than a monolayer is formed, but rarely more than two atomic layers. However, even in such cases, the process is self-limiting.

(iii) Setting the potential anywhere in the UPD region determines the fraction of the surface covered. Thus, an adsorption isotherm can be calculated, relating the fractional surface coverage to the applied potential.

(iv) An additional feature, not shown in Figure 11.1, has been observed experimentally. If the concentration of the metal ion in solution is very low (in the micromole range), deposition becomes mass-transport limited. Under such conditions, even if a negative overpotential is applied, bulk deposition does not start until the formation of a full UPD layer has been completed.

(v) Formation of a uniform UPD layer is independent of the geometry of the part being deposited. This follows from the property of such layers of being *self-limited*. In the language of electroplating, this is referred to as an ideal *throwing power* (see Chapter 18). In electroplating, this would indicate uniform thickness of the coating, independent of the shape of the part being coated. In the case of UPD, each occupied site on the surface of the metal becomes totally inactive (since the potential is positive with respect to the reversible potential), and a uniform UPD layer will be formed everywhere, on

1 This is just a representative number. The exact number varies from metal to metal, and even from one crystal face of the same metal to the other.

crevices as well as on protrusions, around the corner or even on the posterior side of an electrode. How can a single layer of atoms block the surface so efficiently? It is not because it forms a perfect barrier layer. It is because the potential is positive with respect to the reversible potential for bulk deposition of the metal. Formation of further layers is prevented by thermodynamic considerations, not by the ability of a single layer to act as an ideal barrier layer.

(vi) Last but not least, the formation of a UPD layer is a prime example of an adsorption pseudocapacitance. This is discussed in Section 10.2, and the technique of cyclic voltammetry commonly used to study it is discussed in Section 14.4. When a potential is applied, a transient current will flow, which will decay to zero if the potential is held constant. This is exactly the behavior of a capacitor.

The equation controlling the current–potential relationships shown in Figure 11.1 is

$$j_\phi = C_\phi \left(\frac{\mathrm{d}E}{\mathrm{d}t} \right) = C_\phi \nu \tag{11.7}$$

The current is proportional to the sweep rate, and its variation with potential represents the dependence of the adsorption pseudocapacitance on the potential, as discussed in Section 10.2. Although the current observed increases linearly with the sweep rate, the total charge needed to form a UPD layer is independent of it.

If the kinetics of deposition and stripping is fast and there is no mass-transport limitation, the system behaves reversibly, and the cathodic and anodic peaks are observed at the same potential, independent of the sweep rate and of the direction of the sweep. Upon increasing the sweep rate beyond a certain limit, which is characteristic to each system, the cathodic peak starts shifting in the negative direction, while the anodic peak shifts in the positive direction. The system is no longer reversible, and the difference in the peak potentials, ΔE_p, increases with increasing sweep rate. For example, the curves shown in Figure 11.1 indicate that a small degree of irreversibility exists in the case of formation of a UPD layer of Pb on Ag.

UPD of metals is a commonly observed phenomenon, which has been found to occur for dozens of metal couples in both aqueous and nonaqueous solvents. The potential difference ΔE_UPD is independent of the concentration of the metal ion in solution (as long as the UPD layer is formed reversibly and mass-transport limitation is not involved), since both E_p and ΔE_rev follow the Nernst equation, albeit with different values of the standard potential.

The values of ΔE_UPD observed experimentally have been correlated to the difference in the electronic work function of any two metals concerned. Much research has been devoted to the study of UPD on single-crystal metal substrates. As may be expected, ΔE_UPD is different for different crystal planes of the same metal, as are the work function and the potential of zero charge.

11.1.2 UPD on Single Crystals

The mercury/electrolyte interface played a major role in the early studies of the structure of the metal/solution interphase, and electrode kinetics in general. The surface of the liquid metal is highly reproducible, and the low catalytic activity of Hg toward hydrogen evolution provided a rather wide range of potentials where the thermodynamic properties of the interphase could be determined experimentally, allowing theories to be verified or rejected. However, mercury is of little industrial interest, and its use has been all but eliminated in recent decades, because of its high toxicity and its devastating influence on the environment.

The next logical step would have been to study single-crystal electrodes, which represent well-defined surfaces. This is not the way the field developed historically, because it took a long time before high-quality single crystals of metals could be either prepared in the laboratory or purchased commercially; but we need not be concerned with that here. The use of single-crystal substrate has now become common in fundamental studies of UPD for four decades or more.

After a single crystal is prepared and its axis of symmetry is determined, it is cut at a chosen angle with respect to this axis, and embedded in an insulator, so that the solution will be in contact only with a single crystal face. An ideal crystal cut exactly at the correct angle would expose an atomically flat surface to the solution. An error in the angle of cutting will create terraces with atomic steps. With the introduction of scanning tunneling microscopy (STM) to electrochemistry in the 1980s, it became possible to measure the size of such terraces, which were found to be large on the atomic scale, extending to hundreds of nanometers. As a rule of thumb, the larger the error in the angle of cutting, the shorter will the terraces be. Although single crystals may not be atomically flat on the macroscopic scale, the roughness caused by the terraces is not very significant. In any case, the shapes of the cyclic voltammograms, in particular the potential of the peak currents observed for UPD formation, are characteristic of each crystal face. Data for the formation of UPD layers of Pb on the three low-index faces of single-crystal gold are shown in Figure 11.2. On the (111) crystal plane (Figure 11.2a), a very sharp peak is observed. Ideally, this should be the only peak, but there are two smaller peaks, characteristic of two other crystal faces. In Figure 11.2b,c, two major peaks are shown, one of which seems to be associated with the (111) crystal plane, while the others are characteristic of the (100) and (110) crystal planes, respectively. We shall not discuss the details of these peaks, which are only shown here in order to emphasize that different crystal faces give rise to different current/potential curves, as discussed earlier.

Often, the peak current observed for formation of a UPD layer is found to be very narrow, as shown in Figure 11.2a, and this is generally assumed to represent two-dimensional (2D) phase formation. The formation of a 2D phase is controlled by two opposing Gibbs energy terms. On the one hand, we know that UPD must involve the formation of a chemical bond between atoms on the surface of the metal substrate and those in the UPD layer. This would lead to epitaxial (or heteroepitaxial) deposition, which means that the structure of the deposited layer of atoms follows that of the particular crystal plane exposed to the solution. However, this may not be the same as the most stable crystal

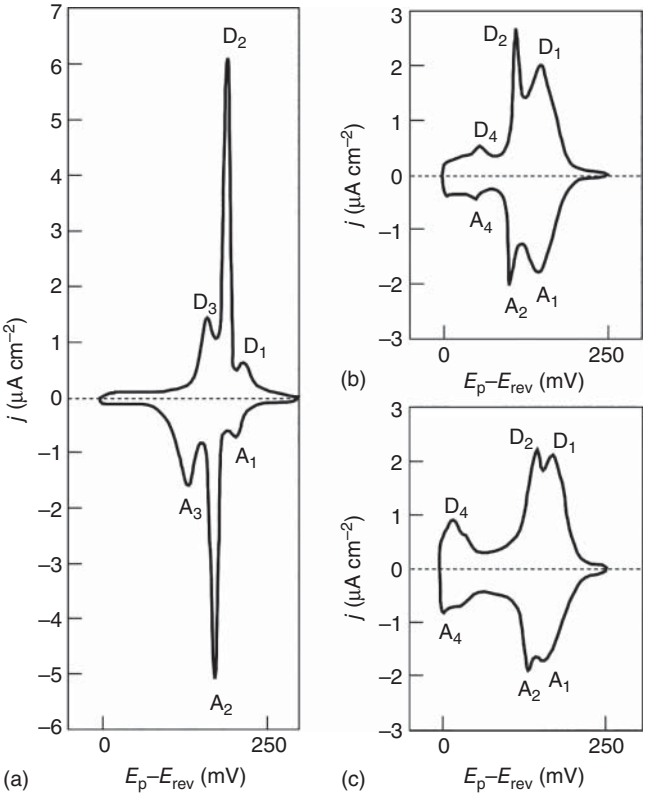


Figure 11.2 Cyclic voltammograms for UPD formation of Pb on different planes of single-crystal Au. (a) (111), (b) (100), (c) (110). $v = 0.42$ mV s^{-1}, 0.5 mM $Pb(ClO_4)_2 + 0.5$ M $NaClO_4 + 5$ mM $HClO_4$. Source: Reproduced with permission from E.B. Budevsky, G.T. Staikov, and J. Lorenz, *Electrochemical Phase Formation and Growth: An Introduction to the Initial Stages of Metal Deposition*, 2008.

structure of the metal deposited in the bulk form. Thus, lead on gold is much less likely to form a 2D phase than silver on gold, because the ratio of crystal radii is only 1.002 for the Au/Ag couple, but 1.21 for the Pb/Au couple.

Formation of a 2D phase on polycrystalline metal surfaces is rarely, if ever, observed, because the metal surface is not atomically flat. Crystallites could be at different angles with respect to each other, and different crystal planes may be exposed to the electrolyte having random orientations, preventing long-range order. Nevertheless, in some cases the surface of a polycrystalline metal is found to behave as if only a single-crystal face was exposed to the solution. The extent of such behavior depends on the method of preparation and pretreatment of the substrate metal. For example, the crystals used for the electrochemical quartz crystal microbalance (see Chapter 16) often have a layer of gold sputtered on the surface, on top of a thin layer of titanium that serves as a binder between gold and the quartz. This gold surface has a (111) preferred orientation, although it is certainly not a single crystal in the usual sense.

11.1.3 Underpotential Deposition of Atomic Oxygen and Hydrogen

It was mentioned earlier, in the discussion of hydrogen evolution on Pt, that a layer of adsorbed hydrogen atoms is formed on the surface, even before the reversible potential for hydrogen evolution has been reached. This could be considered as the formation of a UPD layer of hydrogen. The phenomenon is best seen in cyclic voltammetry, as peak currents appearing at a characteristic potential, positive with respect to the reversible hydrogen electrode in the same solution, as shown in Figure 11.3. The detailed behavior depends on the type of electrode used. Figure 11.3a is the typical curve obtained on a polycrystalline sample (e.g. a wire or a foil). Figure 11.3b was obtained on a spherical single crystal (formed by melting the lower part of a Pt wire), on which different crystal orientations are exposed to the solution. The different peaks represent different energies of adsorption of atomic hydrogen. The area under all the peaks combined shows a charge of about 0.23 mC cm^{-2} of real surface area, namely to a monolayer of adsorbed hydrogen atoms. All the peaks are at potentials positive with respect to the reversible hydrogen electrode.

Cyclic voltammetry on noble metals has been studied extensively. The shape of the curve shown in Figure 11.3a, which is characteristic of Pt in sulfuric acid, is often used as a test for the purity of the system.

The peaks for adsorption and desorption of atomic oxygen are also shown in Figure 11.3a. As with hydrogen, formation of adsorbed oxygen atoms precedes oxygen evolution. The area under the peak corresponds to deposition of a monolayer of oxygen atoms, requiring two electrons per atom deposited. Oxygen evolution does not take place on the bare platinum surface, but on a surface

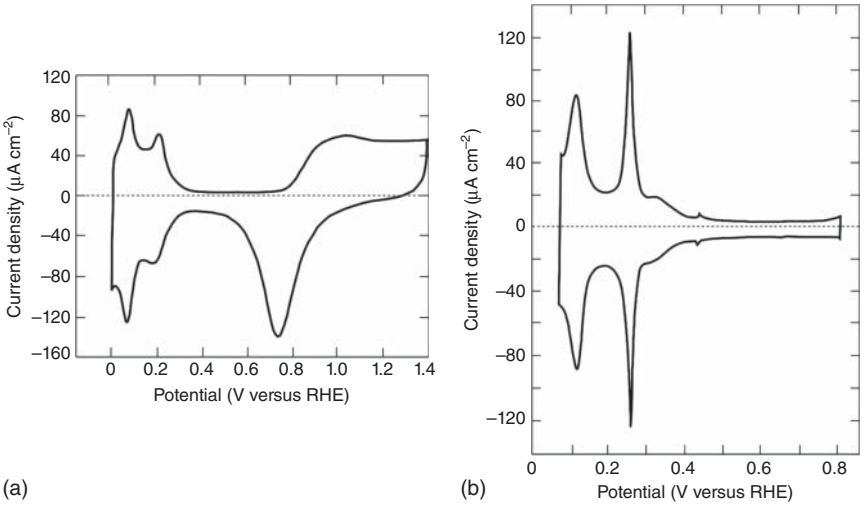


Figure 11.3 Cyclic voltammetry on Pt in 0.5 M H_2SO_4. (a) Polycrystalline Pt, $v = 50$ mV s^{-1}. (b) Spherical single crystal, $v = 50$ mV s^{-1}. The scale of potentials is different because (a) shows also the region of formation and removal of a layer of atomic oxygen, while (b) is limited to the region of adsorption and removal of hydrogen in the double-layer region. Source: Reprinted with permission from J. Clavilier and D. Armand, *J. Electroanal. Chem.* **199** (1986) 187.

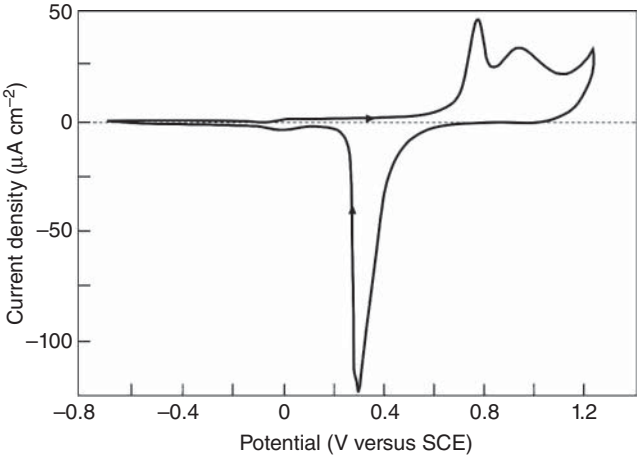


Figure 11.4 Cyclic voltammetry on a single-crystal Au(210) electrode in 10 mM NaF. $v = 20$ mV s^{-1}. Source: Reprinted with permission from A. Hamelin, *J. Electroanal. Chem.* **138** (1982) 395.

modified by a layer of adsorbed oxygen atoms. On the other hand, reduction of molecular oxygen takes place at much more negative potentials and could occur, at least in part, on a bare platinum surface.

The cyclic voltammogram on a gold electrode in the same solution is seen in Figure 11.4. The UPD layer of oxygen atoms is there, but underpotential deposition of hydrogen cannot be detected on gold even at a rather high overpotential. The same is true for mercury, lead, and other soft metals. It would seem that a low exchange current density is associated with the reaction taking place on a bare metal, whereas a high value of j_0 is found when the surface is modified by a layer of adsorbed hydrogen atoms. On the other hand, oxygen evolution does not occur on the bare surface of any metal electrode.

We conclude this Section by noting that UPD is a rather general phenomenon, occurring in both cathodic and anodic reactions. The surface is modified by the UPD layer, and its catalytic activity is altered, usually for the better. The UPD layer is "transparent" to electrons, and probably should be considered to be an extension of the metal, rather than a superficial layer of a foreign substance. A satisfactory understanding of the way in which the surface is modified by the UPD layer, which is essential for the understanding of electrocatalysis, is still lacking.

Further Reading

1 Oviedo, O.A., Reinaudi, L., García, S.G., and Leiva, E.P.M. (2016). *Underpotential Deposition: From Fundamentals and Theory to Applications at the Nanoscale*. Switzerland: Springer International Publishing.

2 Aaronson, B.D.B., Chen, C.H., Li, H. et al. (2013). Pseudo-single-crystal electrochemistry on polycrystalline electrodes: visualizing activity at grains

and grain boundaries on platinum for the Fe^{2+}/Fe^{3+} redox reaction. *J. Am. Chem. Soc.* 135: 3873–3880.

3 Climent, V. and Feliu, J.M. (2011). Thirty years of platinum single crystal electrochemistry. *J. Solid State Electrochem.* 15: 1297–1315.

4 Budevsky, E.B., Staikov, G.T., and Lorenz, J. (2008). *Electrochemical Phase Formation and Growth: An Introduction to the Initial Stages of Metal Deposition*. Weinheim, Germany: Wiley-VCH.

5 Conway, B.E. and Jerkiewicz, G. (2000). Relation of energies and coverages of underpotential and overpotential deposited H at Pt and other metals to the 'volcano curve' for cathodic H_2 evolution kinetics. *Electrochim. Acta* 45: 4075–4083.

6 Szabó, S. (1991). Underpotential deposition of metals on foreign metal substrates. *Int. Rev. Phys. Chem.* 10: 207–248.

7 Gewirth, A.A. (1991). Atomic resolution electrochemistry of underpotential deposition processes. *AIP Conf. Proc.* 241: 253.

8 Clavilier, J. and Armand, D. (1986). Electrochemical induction of changes in the distribution of the hydrogen adsorption states on Pt (100) and Pt (111) surfaces in contact with sulphuric acid solution. *J. Electroanal. Chem.* 199: 187–200.

9 Hamelin, A. (1982). Study of the (210) face of gold in aqueous solutions. *J. Electroanal. Chem.* 138: 395–400.

10 Kolb, D.M., Przasnyski, M., and Gerischer, H. (1974). Underpotential deposition of metals and work function differences. *J. Electroanal. Chem. Interf. Electrochem.* 54: 25–38.

12

Electrosorption

12.1 Phenomenology

12.1.1 What is Electrosorption?

Electrosorption (which is short for "electrochemical adsorption") is a replacement reaction. We have already discussed the role of the solvent at the interphase, in the context of its effect on the double-layer capacitance. It is most important for our present discussion to know that the electrode is always solvated, and the solvent molecules can be adsorbed on the surface both by electrostatic forces and by chemical bonds. Adsorption of a molecule on such a surface requires the removal of an appropriate number of solvent molecules, to make place for the new occupant, so to speak. This is *electrosorption*. In this chapter we shall restrict our discussion to the electrosorption of neutral organic molecules from aqueous solutions, without explicit charge transfer.[1] Using the notation RH for an unspecified neutral organic molecule, we can then represent electrosorption in general by the reaction

$$RH_{soln} + n(H_2O)_{ads} \rightleftharpoons RH_{ads} + n(H_2O)_{sol} \tag{12.1}$$

Several important features of electrosorption follow from this simple equation. First, it becomes clear that the thermodynamics of electrosorption depends not only on the properties of the organic molecule and its interactions with the surface, but also on the properties of water. In other words, the Gibbs free energy of electrosorption is the difference between the Gibbs free energies of adsorption of RH and that of n water molecules:

$$\Delta G_{ads} = \left(\Delta G_{RH_{ads}} - \Delta G_{RH_{sol}}\right) - n\left(\Delta G_{W_{ads}} - \Delta G_{w_{sol}}\right) \tag{12.2}$$

The same relationships also apply to the enthalpy and the entropy of electrosorption. The enthalpy of electrosorption turns out to be less (in absolute value) than the enthalpy of chemisorption of the same molecule on the same surface from the gas phase. On the other hand, the entropy of chemisorption from the gas phase is, as a rule, negative since the molecule RH is transferred from the gas

1 It is recognized here that a species in contact with the surface could interact with the free electrons of the metal, leading to partial charge transfer and imparting an effective charge on the adsorbed species. We distinguish here between this and an explicit charge transfer, which would be written as $RH_{soln} + n(H_2O) \rightarrow R^-_{ads} + H^+ + n(H_2O)_{soln}$.

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

phase to the surface, losing in the process 3 degrees of freedom of translation. This is also true for electrosorption, but in this case n molecules of water are transferred from the surface to the solution, leading to a net increase of $3 \times (n-1)$ degrees of freedom. As a result, the entropy of electrosorption is usually positive. Remembering the well-known thermodynamic relationship

$$\Delta G_{\text{ads}} = \Delta H_{\text{ads}} - T\Delta S_{\text{ads}} \tag{12.3}$$

we conclude that in chemisorption from the gas phase, a negative value of ΔG_{ads} is a result of a negative value of ΔH_{ads}, while the entropy term tends to drive the Gibbs free energy in the positive direction. In electrosorption, the enthalpy term can be less negative or even positive, and it is the positive value of the entropy of electrosorption that renders the Gibbs free energy negative, in most cases. It can be said that *electrosorption is mostly entropy driven, whereas chemisorption is mostly enthalpy driven*.

While the above conclusion is intellectually intriguing, it may also have some important practical consequences, particularly in the area of fuel cells and organic synthesis. Thus, the common wisdom is that the extent of chemisorption *decreases* with *increasing* temperature. This follows formally from the well-known equation

$$\frac{\mathrm{d}\log K_{\text{ads}}}{\mathrm{d}(1/T)} = -\frac{\Delta H^0_{\text{ads}}}{2.3R} \tag{12.4}$$

with negative values of ΔH_{ads} for chemisorption. In contrast, studies of the electrosorption of ethylene on platinum electrodes from acid solutions yielded an enthalpy of adsorption close to zero. The equilibrium constant for the electrosorption of this compound was found to increase with increasing temperature. Thus, raising the operating temperature of a fuel cell, employing ethylene as the fuel, in order to enhance the reaction rate, does not necessarily lower the extent of adsorption of the reactants or intermediates.

The second point to note is that electrosorption depends on the size of the molecule being adsorbed, vis-à-vis its dependence on the number of water molecules that have to be replaced for each RH molecule adsorbed. One may be led to think, on the basis of Eq. (12.2), that large molecules cannot be electrosorbed. This is not necessarily true, because both terms on the right-hand side of Eq. (12.2) may increase with increasing size of the molecule (i.e. with the parameter n), though not necessarily at the same rate.

An additional unique feature of electrosorption is that the coverage is a function of potential, at constant concentration in solution. Thus, we can discuss two types of isotherms: (i) those yielding θ as a function of c_{b}, and (ii) those describing the dependence of θ on E. This *is not* a result of faradaic charge transfer. Neither is it due to electrostatic interactions of the adsorbed species with the field inside the compact part of the double layer, considering that a potential dependence is observed even for neutral organic species having no permanent dipole moment. As we shall see, it turns out that the potential dependence of θ can be caused by the dependence of the Gibbs free energy of adsorption of water molecules on potential.

12.1.2 Electrosorption of Neutral Organic Molecules

Electrosorption has been studied on mercury more than on any other metal, not because this is the most interesting system, either from the fundamental or the practical point of view, but because it is the easiest system to study, and because the results obtained are not complicated by uncertainties resulting from different features of the surface – a problem common to the study of solid surfaces. The dependence of θ on potential for the adsorption of *n*-butanol on mercury was shown in Figure 9.8. In Figure 12.1 we show plots of the fractional surface coverage, θ, for the electrosorption of phenol with methanol or water as a solvent. One should note that the dependence of θ on E is roughly bell-shaped in this case, and

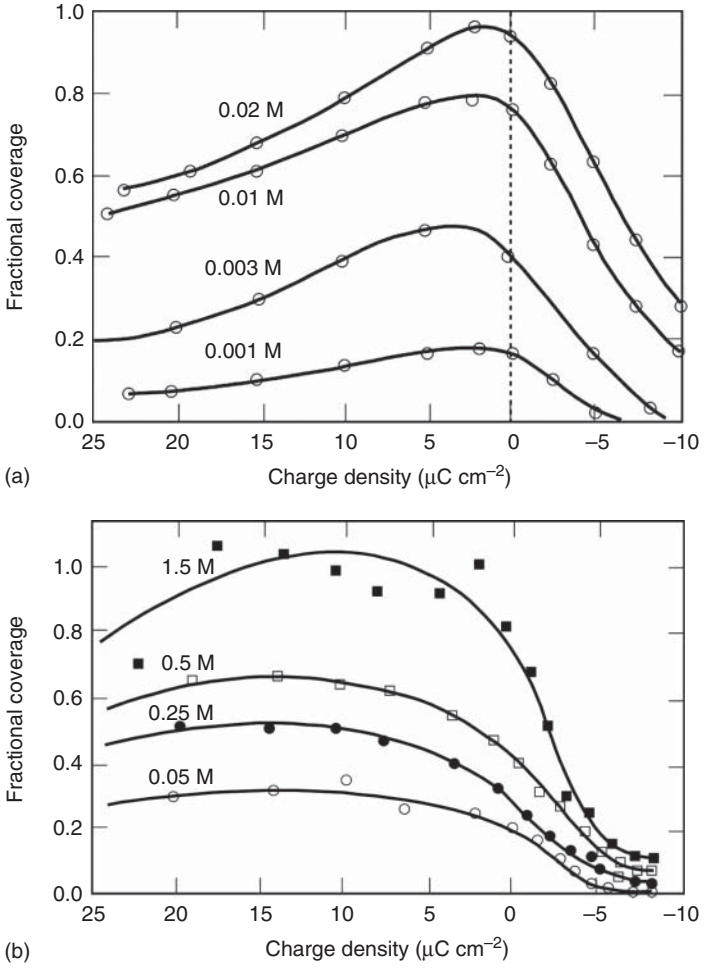


Figure 12.1 The electrosorption of phenol on mercury, from (a) water, and (b) methanol as the solvent, as a function of the charge density. Supporting electrolyte: 0.1 M LiCl. The concentrations of phenol are marked on the curves. Source: Based on data by J.O'M. Bockris, E. Gileadi, and K. Müller, *Electrochim. Acta* **12** (1967) 1301.

also for many other neutral molecules, whether they do or do not have a permanent dipole moment. Typically, the maximum of adsorption occurs at a potential that is slightly negative with respect to the potential of zero charge.

If electrosorption is restricted to a monolayer, the fractional surface coverage can be related to the surface excess simply by writing

$$\theta \equiv \frac{\Gamma}{\Gamma_{\max}} \tag{6.9}$$

This is a satisfactory approximation in the present as well as in similar cases, since the interaction of a neutral organic molecule with the surface does not extend beyond the thickness of one monolayer.

Comparing the data in water and in methanol, we note that a similar extent of coverage is reached in water in the range of concentrations of 1–20 mM, and in methanol in the range of 0.05–1.5 M. Considering that it is the same molecule being adsorbed on the same surface in both cases, the difference must be associated mainly with the difference in the solubility of phenol in the two solvents. This leads us to discuss the question of the appropriate scale of concentration to be used when comparing isotherms measured in different solvents. Usually, chemists prefer to express concentrations in units of moles per liter (or moles per cubic decimeter). This is fine for aqueous solutions (or, at least, for a fixed solvent), but it fails totally when different solvents are compared. Similarly, when comparing the adsorption of different solute molecules in the same solvent, the solubility of the adsorbed molecule also plays an important role. In such cases, it is best to use a dimensionless scale of the bulk concentration divided by the saturation concentration in the same solvent, c_b/c_{sat}. This follows directly from the fact that the chemical potential in a saturated solution is always equal to that of the solute in a pure solid phase, irrespective of either the solute or the solvent. Consequently, the choice of the above dimensionless scale of concentration permits us to compensate for the differences in the Gibbs free energy of interaction between the solvent and the solute in the bulk of the solution, and the effects seen arise only from the different interactions of the solutes with the surface. This is a very good way of comparing the adsorption of *different solutes* from the *same solvent*. When the adsorption of the *same solute* from *different solvents* is studied, the choice of this dimensionless concentration still ignores the different interaction of the solvent molecules with the surface, as given by Eq. (12.2), so it is an approximation, albeit a much better one than using molar concentrations.

12.1.3 The Potential of Zero Charge, E_{pzc}, and its Importance in Electrosorption

The concept of the potential of zero charge, E_{pzc}, has already been discussed in the context of electrocapillary thermodynamics, where we showed that, for an ideally polarizable interphase, E_{pzc} coincides with the electrocapillary maximum. In view of the very high accuracy attainable with the electrocapillary electrometer, it is possible to determine the value of E_{pzc} for liquid metals near room temperature to within about ± 1 mV.

Many methods have been used to determine the value of E_{pzc} on solid electrodes. The one that seems to be most reliable, and relatively easy to perform, is based on the diffuse double-layer theory. Measurement of the capacitance in dilute solutions ($c_b \leq 0.01$ M) should show a minimum at E_{pzc}, as seen in Eq. (8.10) and Figure 8.4. Lowering the concentration yields better defined minima. Modern instrumentation allows us to extend the measurement of capacitance to low concentrations of the electrolyte (see Section 15.1), increasing the accuracy of the determination of E_{pzc} on solid electrodes. One should bear in mind, however, that the minimum in capacitance coincides with E_{pzc} only if a symmetrical electrolyte, such as NaF, is used. Asymmetric electrolytes of the type A_mB_n (where $m \neq n$) also show minima in the plot of capacitance versus potential in dilute solutions, but the minima are shifted slightly from E_{pzc}. The symbol $\overline{E}$-E represents the potential measured with respect to the potential of zero charge.

Although the instrumental aspects of measuring E_{pzc} can yield quite accurate results, chemistry is lagging behind. The most difficult problem, as always with solid electrodes, is the lack of reproducibility of surface preparation. The most reliable results have been obtained on single crystals of noble metals, where E_{pzc} depends on the particular crystal face (plane) exposed to the solution. An exceptional example of this is shown in Figure 12.2, where E_{pzc} is plotted for a large number of crystal faces of gold. We need not go into the details of these different crystal planes. Figure 12.2 is used just to illustrate that E_{pzc} can be measured

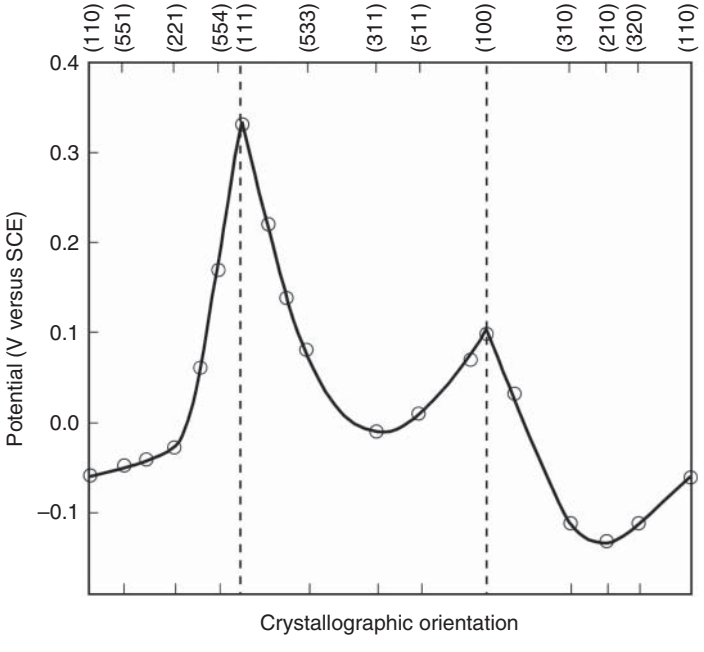


Figure 12.2 The potential of zero charge for single-crystal gold, plotted as a function of crystal orientation. Source: Reproduced by permission from J. Lecoeur, J. Andro, and R. Parsons, *Surf. Sci.* **114** (1982) 320.

rather accurately, even on solids, and that it is clearly a function of the crystallographic orientation. It should also be noted that the difference in the values of E_{pzc} for different crystal planes can be quite large, amounting to about 0.4 V between the {110} and the {111} families of planes of gold, for example.

On a polycrystalline sample certain crystal faces may dominate, depending on the mechanical, thermal, and electrochemical pretreatment of the sample, giving rise to different values of the E_{pzc}. It was shown, for example, that cycling the potential of a platinum electrode between oxygen and hydrogen evolution for a long time causes faceting, with the (111) crystal face becoming predominant.

Determination of E_{pzc} on base metals such as copper, nickel, or iron can be complicated by the formation of oxide layers. The value of E_{pzc} measured may then correspond to an oxide-covered surface, rather than to the bare metal.

The occurrence of faradaic reactions of any kind, and particularly those leading to the formation of adsorbed intermediates, can severely interfere with the determination of E_{pzc}, when based on the measurement of the capacitance minimum. The high values of the adsorption pseudocapacitance, C_{ϕ}, which extends over a significant range of potential, can distort the measurement of double-layer capacitance in dilute solutions, as discussed in Section 10.2.3.

Values of E_{pzc} are shown in Table 12.1, for a number of solid metals. The value for mercury is also included, for comparison. Electrosorption depends primarily on the excess charge density, q_M, which is related, albeit not quite linearly, to the potential measured with respect to the potential of zero charge.

Coverage by neutral organic species decreases at both negative and at positive values of the potential (or the charge), with a maximum of coverage at about $q_M \approx -2\ \mu C\,cm^{-2}$. On the scale of potential, the region of significant coverage extends over about 0.8 V, $\overline{E} \approx (-0.6\ \ to + 0.2)\ \ V$. For aromatic compounds, the

Table 12.1 The potential of zero charge, E_{pzc}, for different metals versus SHE.

Metal	E_{pzc} (V)	Electrolyte	c_b (mM)
Cd	−0.75	KCl	1
Tl	−0.71	NaF	1
Ga	−0.69	HCl	1
Zn	−0.63	HCl	1
Pb	−0.56	NaF	1
Bi	−0.39	HCl	1
Sn	−0.38	K_2SO_4	2
Fe	−0.37	H_2SO_4	1
Hg	−0.193	NaF	1
Ag(111)	−0.46	NaF	1
Au(111)	+0.19	NaF	5
Pt	+0.18	HF	300

coverage declines more slowly on the positive side, probably due to interaction of the π-electrons with the metal. For charged species, the situation is more straightforward. Positively charged ions are adsorbed mostly at negative values of $\overline{E}$, and vice versa. This effect is superimposed on other factors controlling electrosorption, so that a negatively charged molecule may be specifically adsorbed to some extent even on a negatively charged surface, as a result of the chemical energy of interaction between the molecule and the surface. Naturally, the dependence of θ on potential will not be symmetrical in this case as it is, more or less, for neutral molecules.

It should be noted here that the adsorption of intermediates formed by charge transfer is not controlled by the potential of zero charge. When a process such as

$$H_3O^+ + e_M^- + Pt^0 \rightleftharpoons Pt\text{-}H + H_2O \tag{10.5}$$

which leads to an adsorption isotherm of the form

$$\frac{\theta}{1-\theta} \exp\left(\frac{r}{RT}\theta\right) = K_0 c_b \quad \exp\left(\frac{F}{RT}E\right) \tag{10.13}$$

the region of potential over which θ is significant depends on the equilibrium constant K_0, which is related to the Gibbs free energy of adsorption. The value of K_0 depends on the metal through its dependence on the Gibbs free energy of the M—H bond, but it is not directly dependent on E_{pzc}.

12.1.4 The *Work Function* and the Potential of Zero Charge

The *work function* Φ is the energy required to remove an electron from the bulk of a metal to infinity in vacuum. It is an intrinsic property of the metal, which is measured under conditions of ultrahigh vacuum. This is measured on surfaces that have been meticulously cleaned. Even a small amount of impurity adsorbed on the surface (e.g. oxygen or an oxide, water, or carbonaceous molecule) could give rise to significant errors in the measured value of the work function. Moreover, the value measured on a polycrystalline metal surface is a weighted average of the work function values for different crystal faces. Thus, each metal has, in fact, several work function values, each characteristic of a different crystal face. The relevant scientific literature is replete with data with values of Φ for different metals and many of the common crystal faces for each metal.

The electrochemical properties of the metal/electrolyte interphase follow a similar pattern, but involve some complication. Thus, the value of E_{pzc} is found to be different for different crystal faces, and so are the cyclic voltammograms characteristic to UPD formation, as shown in Chapter 11. It has been suggested that E_{pzc} and Φ should be linearly related to each other, and indeed such correlations have been reported in the literature. However, while Φ is a characteristic property of each metal and each crystal face, E_{pzc} is a property of the metal/electrolyte interphase, not just the metal. This is not surprising, in view of the fact that an electrode in contact with an electrolyte is never clean, on the level of the cleanliness of a surface in ultrahigh vacuum. In the best case, it is covered by pure solvent molecules and influenced by the ions of the ubiquitously present solute, needed to impart electrolytic conductivity. In addition, catalytic

metals, which are of main interest in technology, may have adsorbed hydrogen or some oxygen-containing species, or formed surface oxide.

12.2 Adsorption Isotherms for Neutral Species

12.2.1 General Comments

An isotherm describing electrosorption can be written in general form as

$$f(\theta) = Kc_b g(E) \tag{12.5}$$

where $f(\theta)$ and $g(E)$ represent some, as yet unspecified, functions of the fractional surface coverage and the potential, respectively. Up to now we have been concerned with the form of the function $f(\theta)$. The potential dependence of θ for an adsorbed species formed by charge transfer had the form of the Nernst equation

$$E = -\frac{2.3RT}{F}\log K_1 + \frac{2.3RT}{F}\log\left[\left(\frac{1}{c_{b,Cl^-}}\right)\left(\frac{\theta}{1-\theta}\right)\right] \tag{6.25}$$

Here, we shall discuss the dependence of coverage on the potential in the absence of charge transfer, paying particular attention to the size of the adsorbed molecules, expressed in terms of the number of sites, n, on the surface occupied by each molecule, which is equal to the number of water molecules replaced from the surface per molecule of the adsorbed species.

Historically, one of the early triumphs of the Langmuir isotherm was in distinguishing between physically adsorbed molecular hydrogen and chemically adsorbed atomic hydrogen. For the former, one writes

$$H_2 + M \rightleftharpoons M{-}H_2 \tag{12.6}$$

which should follow the Langmuir isotherm,

$$\frac{\theta}{1-\theta} = K_1 p_{H_2} \tag{12.7}$$

On the other hand, for the adsorption of atomic hydrogen the appropriate equilibrium is

$$H_2 + 2M \rightleftharpoons 2M{-}H \tag{12.8}$$

In the second case, the rate of adsorption is proportional (approximately) to $(1-\theta)^2$, and the rate of desorption is proportional to θ^2, since we consider empty and occupied sites to be reactants for the adsorption and the desorption reactions, respectively. The corresponding isotherm is

$$\left[\frac{\theta}{1-\theta}\right]^2 = K_2 p_{H_2} \tag{12.9}$$

Thus, at low coverage, dissociative adsorption leads to

$$\theta = \sqrt{K_2 p_{H_2}} \tag{12.10}$$

where the subscripts in K_1 and K_2 are added only to show that the equilibrium constants in Eqs. (12.7) and (12.9) have different numerical values.

When a larger molecule taking up n sites on the surface is being adsorbed, we could follow the same line of reasoning and write

$$\frac{\theta}{(1-\theta)^n} = Kp \quad (or = Kc_b) \tag{12.11}$$

This is not a rigorously correct equation, but it is a better approximation than the Langmuir isotherm to describe the adsorption of larger molecules.

For the electrosorption of a large organic molecule, represented by the equilibrium

$$RH_{soln} + n(H_2O)_{ads} \rightleftharpoons RH_{ads} + n(H_2O)_{sol} \tag{12.1}$$

a better approximation can be derived for the function $f(\theta)$ in Eq. (12.12) by expressing the equilibrium constant in terms of the mole fractions of the species involved.

$$K = \frac{(X_{RH,ads})(X_{W,soln})^n}{(X_{RH,soln})(X_{W,ads})^n} \tag{12.12}$$

Substituting the appropriate values for the mole fraction of the adsorbed molecule in the solution and on the surface, we obtain the expression

$$Kc_{b,RH} = \left(\frac{\theta}{(1-\theta)^n}\right)\left(\frac{\left[\theta + n(1-\theta)^{n-1}\right]}{n^n}\right) \tag{12.13}$$

It is easy to see that if we set $n = 1$ in Eq. (12.13), it reverts to the Langmuir isotherm. This should not be surprising, because no coverage-dependent Gibbs free energy of electrosorption has been introduced in deriving Eq. (12.13).

12.2.2 The Parallel-Plate Model of Frumkin et al.

We now turn to the potential dependence of electrosorption of neutral molecules, considering first the parallel-plate model. This is a phenomenological model, which depends on considerations of the changes in the electrostatic energy of the interphase caused by adsorption. Assuming that measurements are taken in concentrated solutions of a supporting electrolyte, we can neglect diffuse double-layer effects and focus our attention on the Helmholtz part of the double layer, considered as a parallel-plate capacitor. In a pure electrolyte, the capacitance C_0 and the corresponding surface charge density, q_0, are determined by the properties of water at the interphase, mainly its effective dielectric constant and its dimensions (which determine the thickness of the capacitor). We have already seen that adsorption of an organic molecule tends to decrease the capacitance. This property can be associated with the combined effects of a lower dielectric constant and an increase in the thickness of the parallel-plate capacitor. Writing q_0 and q_1 for the charge densities at $\theta = 0$ and $\theta = 1$, respectively, the equation proposed is

$$q_0 = C_0\overline{E} \quad \text{and} \quad q_1 = C_1(\overline{E} - E_N) \tag{12.14}$$

where C_0 and C_1 are the values of the double-layer capacitance at $\theta = 0$ and $\theta = 1$, respectively, and $\overline{E} \equiv E - E_{pzc}$ is the potential with respect to the potential of zero

charge. The parameter E_N represents the shift in the potential of zero charge, caused by a full monolayer of adsorbed species.

The Frumkin isotherm for the adsorption of neutral species is based on the assumption that at any value of the coverage, the interphase can be viewed as two capacitors connected in parallel. It follows that the charge q_0 corresponding to a given value of θ can be written as

$$q_\theta = q_0(1-\theta) + q_1\theta = C_0\overline{E}(1-\theta) + C_1(\overline{E} - E_N)\theta \tag{12.15}$$

The derivation of this isotherm for neutral species is rather involved and is not given here. We note only that it is based on calculating the difference in electrostatic energy of charging the double-layer capacitor with and without the adsorbed species. The final result is as follows:

$$K_{\overline{E}} = K_{E_{pzc}} \exp\left(-\frac{0.5(C_0 - C_1)\overline{E}^2 + C_1\overline{E}E_N}{RT\Gamma_{max}}\right) \tag{12.16}$$

where $K_{\overline{E}=0}$ is the value of the equilibrium constant at E_{pzc}. If we ignore the term containing E_N, we find that this isotherm predicts a maximum of adsorption at E_{pzc}, with θ declining symmetrically on either side. The term $C_1\overline{E}E_N$ accounts for the failure of the potential of maximum adsorption observed experimentally to coincide with E_{pzc}.

All the quantities that appear in the exponent of Eq. (12.16) can be measured, at least in principle, thus permitting the isotherm to be tested experimentally. This is a great advantage of any theory, inasmuch as it does not contain any adjustable parameters. The disadvantage of this model is that it makes no attempt to explain the observed phenomena on the molecular level. Thus, the values of $(C_0 - C_1)$ and E_N are taken as such – it is not explained why they have their observed values, or how these values depend on molecular size, on the orientation of the molecules at the interphase, and on the nature of the interactions with the surface.

The assumption underlying the derivation of the above isotherm is tantamount to assuming that the surface charge density is a linear function of the coverage at constant potential, as seen in Eq. (12.15). This is by no means generally correct, although it may be a good approximation in many cases.

Equation (12.14) can be rewritten in the form

$$\theta = \frac{q_\theta - q_0}{q_1 - q_0} \tag{12.17}$$

which may be used to determine the coverage from differential capacitance measurements, if the potential of zero charge is known with sufficient accuracy to allow the determination of the charge by integration of the capacitance as a function of the potential.

An approximate form of Eq. (12.17) is

$$\frac{dq}{dE} = \frac{dq_0}{dE}(1-\theta) + \frac{dq_1}{dE}\theta - (q_0 - q_1)\frac{d\theta}{dE} \tag{12.18}$$

If one neglects the potential dependence of θ in the last term of the above equation, we have

$$\theta \approx \frac{C_\theta - C_0}{C_1 - C_0} \tag{12.19}$$

This may be a reasonable approximation in certain cases, but it must be considered inherently inconsistent, when the purpose of the experiment is to determine the potential dependence of the surface coverage θ.

12.2.3 The Water Replacement Model of Bockris et al.

Another isotherm that has been studied in detail is centered on the role of water at the interphase. As noted earlier, electrosorption is a replacement reaction and its standard Gibbs energy of adsorption is the difference between the standard Gibbs free energies of adsorption of the organic species and of the water molecules it replaces from the surface (see Eqs. (12.1) and (12.2)). Thus, the dependence of θ on the potential is related to the variation of the standard Gibbs free energy of adsorption of water with potential. The electrostatic energy of interaction of the water dipole $\vec{\mu}$ with the electrostatic field $\vec{F}$ in the Helmholtz double layer is given by

$$U_\mu = \vec{\mu}\vec{F}\cos\phi \tag{12.20}$$

A fundamental premise of this isotherm is that water molecules can take up only one of two positions at the interphase – with the dipole vector either in the direction of the field or in the opposite direction. Thus, the energy of interaction between water dipoles and the field is either $\vec{\mu}\vec{F}$ or $\overleftarrow{\mu}\vec{F}$, depending on its orientation.

Considering only the energy of interaction of the water dipole with the surface charge density, all the water molecules should have their positive ends facing the surface when it is negative, and vice versa, with a sharp transition at the E_{pzc}. There are, however, several factors that modify this transition, making it gradual. The following statements can be made:

1. First, there is the entropy term. A situation in which all dipoles are oriented in one direction represents the minimum entropy. Thus, enthalpy and entropy affect the orientation in opposite directions, as is often found in chemistry. A compromise, determined by the lowest Gibbs free energy, is reached. The resulting orientation of water molecules at the interphase depends both on the temperature and the potential. Let us define here a parameter Z as

$$Z = \frac{\vec{N} - \overleftarrow{N}}{N_{\text{T}}} \tag{12.21}$$

where $\vec{N}$ and $\overleftarrow{N}$ are the number of water molecules per unit surface area in the two allowed orientations, depending on charge density, as shown schematically in Figure 12.3.

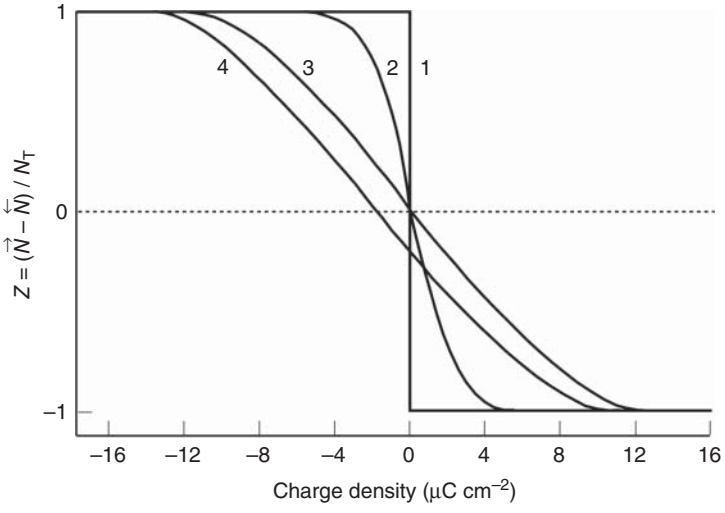


Figure 12.3 Schematic representation of the variation of water orientation at the interphase with charge density. (1) Only the energy of interaction between the dipole and the field in the double layer is considered, (2) an entropy term is added, (3) lateral dipole–dipole interactions are also considered, and (4) the difference in chemical energy of adsorption of water in the two orientations is also taken into account.

2. Next, there is chemical interaction of water with the surface. A water molecule cannot be regarded as a structureless electric dipole. It is more strongly adsorbed when the negative end of the dipole (i.e. the oxygen atom) is facing the surface. Hence, at $q = 0$ the parameter Z is negative. In order to force the parameter Z to be zero, one must apply a small negative potential or charge. In the case of neutral organic molecules on mercury, this amounts to about $q \approx -2$ μC cm^{-2}.
3. Finally, we must consider the lateral interaction term. If left alone (i.e. in the absence of an external field), dipoles would pair up to minimize the total electrochemical Gibbs energy of the system, yielding a value of $Z = 0$. This opposes the tendency of the field to orient all the dipoles in one direction, making the transition from one orientation to the other even slower.

We shall again skip the tedious part of deriving the isotherm and write the final result in a simplified form, as follows:

$$K(E) = K(E_{\text{pzc}}) \exp\left[-nZ\left(\frac{\vec{\mu}\vec{F} - Z\sigma}{kT}\right)\right] \tag{12.22}$$

It is important to consider the physical origin of the various terms in Eq. (12.22). The term nZ is the number of water molecules removed from the surface for each organic molecule adsorbed, multiplied by the fractional excess of water molecules oriented one way or the other. The parameter Z is defined in such a way that the product $\vec{\mu} \times \vec{F}$ is always positive, leading to a symmetrical decrease of the equilibrium constant (and hence, of the coverage) on both sides of $Z = 0$. The parameter σ is the total electrical energy of interaction of each water dipole with

all other dipoles surrounding it. It is also multiplied by Z, since interactions with oppositely oriented dipoles cancel each other.

The complete isotherm has the form

$$\left[\frac{\theta}{(1-\theta)^n}\right]\left[\frac{\theta + n(1-\theta)^{n-1}}{(n^n)}\right] = K(E_{\text{pzc}})c_{\text{b,RH}}\exp\left(-nZ\frac{\vec{\mu}\vec{F} - Z\sigma}{kT}\right) \tag{12.23}$$

Although this may look rather complex, it is a simplified form of the isotherm. Thus, the function $f(\theta)$ on the left-hand side of Eq. (12.23) does not take into account changes in the standard Gibbs free energy of adsorption due to surface heterogeneity or to lateral interactions between the adsorbed molecules. Since the unique feature of this isotherm is the relation between the potential and the surface coverage, we write

$$f(\theta) = K(\overline{E} = 0)c_{\text{b,RH}}\exp\left[-nZ\left(\frac{\vec{\mu}\overline{E}/\delta - Z\sigma}{kT}\right)\right] \tag{12.24}$$

The isotherm shown in Eq. (12.23) can be fitted to the experimental results, and there have been additional attempts to describe the dependence of the fractional surface coverage on solid substrate on the concentration of the species in solution and on the charge at the interphase. The most important point is embedded in the beginning of this chapter (see Eq. (12.1)), namely that the solvent plays an important role in electrosorption, because it is a replacement process. Unlike the adsorption from the gas phase, where the entropy of adsorption is *always negative,* in electrosorption the entropy could go either way, depending on the number of water molecules replaced from the interphase.

Further Reading

1 Su, X. and Hatton, T.A. (2016). Electrosorption. In: *Kirk-Othmer Encyclopedia of Chemical Technology* (ed. J.G. Speight), 1–11. Wiley.

2 Lecoeur, J., Andro, J., and Parsons, R. (1982). The behaviour of water at stepped surfaces of single crystal gold electrodes. *Surf. Sci.* 114: 320–330.

3 Gileadi, E. (ed.) (1967). *Electrosorption*. New York, USA: Plenum Press.

13

Fast Transients, the Time-Dependent Diffusion Equation, and Microelectrodes

13.1 The Need for Fast Transients

13.1.1 General

In this chapter we shall focus our attention on the use of transients for the separation between activation-controlled and mass-transport-controlled processes. The need to enhance the rate of mass transport has already been discussed briefly in Section 4.2.2. It follows from the fact that mass transport and charge transfer are consecutive processes, and therefore the slower of the two will be rate-determining. This is expressed mathematically by

$$\frac{1}{j} = \frac{1}{j_{ac}} + \frac{1}{j_L} \tag{13.1}$$

For values of $j/j_L \ll 1$, the measured current density is close to the activation-controlled current density, and thus the latter can be determined accurately. When the measured current density approaches its limiting value (where j/j_L is in the range of 0.1–0.9), the value of j_{ac} can still be calculated, employing Eq. (13.1), provided that the limiting current density is known; however, the accuracy is reduced as the measured current density approaches mass-transport limitation. This follows from Eq. (13.1), which can also be rewritten as

$$\frac{1}{j_{ac}} = \frac{j_L - j}{j_L \times j} \tag{13.2}$$

For values of the measured current density approaching the limiting current density, the value of j_{ac} calculated from the above equation depends on a small difference between two large numbers, hence its accuracy is diminished. To illustrate this point, assume that the values of j_L and j are 100 and 90 mA cm^{-2}, respectively, and the accuracy in measuring each of these quantities is ± 1 mA cm^{-2}. In this case, the difference, $j_L - j = 10$ mA cm^{-2}, will be determined with an accuracy of about ± 1.4 mA cm^{-2}. Consequently, the error in determining j_{ac} will be about 14%, although the raw measurement of current is accurate to within 1%.

The following two points of interest should be noted in relation to the above calculation:

(a) Using the above values of j_L and j, we find that $j_{ac} = 900$ mA cm^{-2}. This may seem odd. The limiting current, by definition, represents the highest rate

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

at which the reaction can proceed, so how can the activation-controlled current be nine times higher? The answer is given by examining Eq. (13.1) carefully. The activation-controlled current is not the rate at which the reaction occurs. It represents the rate at which *it would have occurred* had there been no mass-transport limitation. When it is much higher than the limiting current, it has very little effect on the observed rate of the electrochemical reaction. In other words, the measured current will be close to the mass-transport-limited current, which is exactly what we have assumed in the above calculation. Thus, one can determine the rate of an electrochemical charge-transfer process (represented by j_{ac}) even when this rate exceeds the rate of mass transport – albeit with significant loss of accuracy. This is shown schematically in Figure 1.1.

(b) Is the random error of 1% assumed above reasonable? A current of 100 mA can be measured with an accuracy of 1 μA, employing equipment available in most electrochemistry research laboratories. However, in our *digital era* the sensitivity and accuracy of such measurements are no longer limited by the quality of the electronic instrumentation – it is determined by the stability of the chemical system. Many factors determine the latter, and we shall not dwell on them here. However, careful experimental practice (such as high purity of the solution, careful elimination of oxygen, good control of temperature, and proper correction for residual background currents) can reduce the random error significantly.

It can be concluded that increasing the rate of mass transport is important in studying electrode kinetics, because it increases the upper limit of current density where the kinetic parameters can be determined. The limiting current density, j_L, can be generally written in the form

$$j_L = \frac{nFDc_b}{\delta} \tag{1.4}$$

where the Nernst diffusion layer thickness, δ, is determined by the mode of mass transport. In quiescent solutions and in the presence of an excess of supporting electrolyte (where convective mass-transport migration in the electrical field gradient is avoided), mass transport is dominated by diffusion. Under these conditions

$$\delta = (\pi Dt)^{1/2} \tag{4.11}$$

and the diffusion-limited current density is given by

$$j_L = \frac{nFDc_b}{\sqrt{\pi Dt}} \tag{13.3}$$

Thus, the diffusion-limited current density is proportional to $t^{-1/2}$, therefore fast transients can extend the range where the activation-controlled current density can be determined.

The equivalent circuit for a system in which diffusion plays a significant role is shown in Figure 13.1. The symbol "W" stands for the *Warburg impedance*, which accounts for mass-transport limitation by diffusion. This is an odd element in the

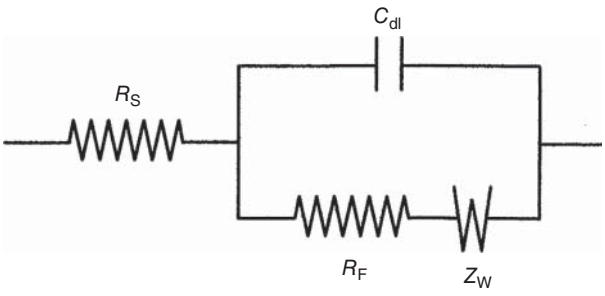


Figure 13.1 The equivalent circuit for partial mass-transport limitation by diffusion. The Warburg impedance, Z_W, is in series with the faradaic resistance, R_F, and in parallel with the double-layer capacitance, C_{dl}.

Table 13.1 Comparison of values of δ obtained by different methods ($n = 2$, $D = 6 \times 10^{-6}$ cm^2 s^{-1}).

Type of stirring	δ (μm)	Corresponding time for transient/s $t = \delta^2/\pi D$	j_L (mA cm^{-2}) ($c_b = 10$ mM)
Natural convection	150–250	12–33	0.80–0.48
Magnetic stirrer	50–100	1.3–5.3	2.4–1.2
RDE at 400 rpm	30	0.48	4.0
RDE at 10^4 rpm	6.0	19×10^{-3}	20
Impinging jet	2.0	2×10^{-3}	60
Fast pulse	0.14–1.4	1×10^{-5}–1×10^{-3}	860–86

equivalent circuit, since it behaves in some ways as a capacitor and in other ways as a resistor, as will be discussed in Section 15.1

The strength of the fast-transient methods in the study of electrode reactions is shown in Table 13.1. The fastest methods of stirring the solution can be matched by transients of a few milliseconds. The limitations of fast transient measurements, both from the experimental and the theoretical points of view, are discussed further.

13.1.2 Small-Amplitude Transients

Transient measurements can be of two types: small-amplitude transients that give rise to a linear response, and large-amplitude transients that result in a nonlinear, often exponential, response. We have already seen (see Section 5.2.4) that a system at equilibrium responds linearly to a small perturbation in potential or in current, according to the equation

$$\frac{j}{j_0} = \frac{n}{\nu}\frac{\eta F}{RT} \tag{13.4}$$

A "small" perturbation in this context is one for which $n\eta F/vRT \ll 1$ (corresponding to $j/j_0 \ll 1$). The linearity of the response allows easier and more rigorous mathematical treatment and is, therefore, often preferred. A small current or voltage perturbation also implies that the changes in concentration of the reactants and products near the electrode surface are small, and the associated equations of mass transport can also be linearized, to simplify the mathematical treatment.

13.1.3 The Sluggish Response of the Electrochemical Interphase

Considering Table 13.1 we note that the use of fast transients in the range of 0.01–1.0 ms yields limiting currents that are higher than those achieved by the fastest methods of stirring. It would be nice to extend this method by the use of very short pulses, in the range of nanoseconds, to study even faster electrode reactions. This is, unfortunately, not possible, because of the sluggishness of the interphase, which is due to the need to charge the double-layer capacitor. Thus, if we wish to change the potential by, say, 10 mV, we need to add a charge of

$$\Delta q = C_{\mathrm{dl}} \times \Delta\eta = 20\ \ \mu\mathrm{F}\ \ \mathrm{cm}^{-2} \times 10\ \ \mathrm{mV} = 200\ \ \mathrm{nC}\ \ \mathrm{cm}^{-2} \tag{13.5}$$

Supplying this amount of charge in 1 ns requires an average current density of 200 A cm^{-2}, which is clearly impractical. Extending the pulse duration to $t = 0.01$ ms requires an average current density of only 0.02 A cm^{-2}, which is well below the diffusion limited current density shown in Table 13.1. In practice, studies with pulses of 0.01 ms duration are relatively easy. Very careful design of the electrochemical cell and its connection to the measuring system can allow measurements at 10^{-6}–10^{-7} seconds, but not at shorter times. This has nothing to do with the limitations of the electronic instruments, of course; it is an inherent physical limitation of the behavior of the metal/solution interphase.

13.1.4 How can the Slow Response of the Interphase be Overcome?

13.1.4.1 Galvanostatic Transients

Imagine that we wish to determine the current–potential relationship for a given reaction, over a wide range. We can apply a series of galvanostatic steps and observe the steady-state potential corresponding to each current density. The trick is to do the measurement rapidly, before diffusion limitation starts to play a role. On the other hand, we must wait long enough for the double layer to be charged up to its steady-state value, when the potential across it is given by jR_{F}. The results of this type of experiment, for different kinetic parameters, are shown in Figure 13.2. There are two parameters in this figure that are characteristic to the interphase – the time constant of the parallel combination of a capacitor and a resistor, τ_{c}, given by

$$\tau_{\mathrm{c}} \equiv C_{\mathrm{dl}} \times R_{\mathrm{F}} = \frac{\nu}{n}\frac{RT}{F}\frac{C_{\mathrm{dl}}}{j_0} \tag{13.6}$$

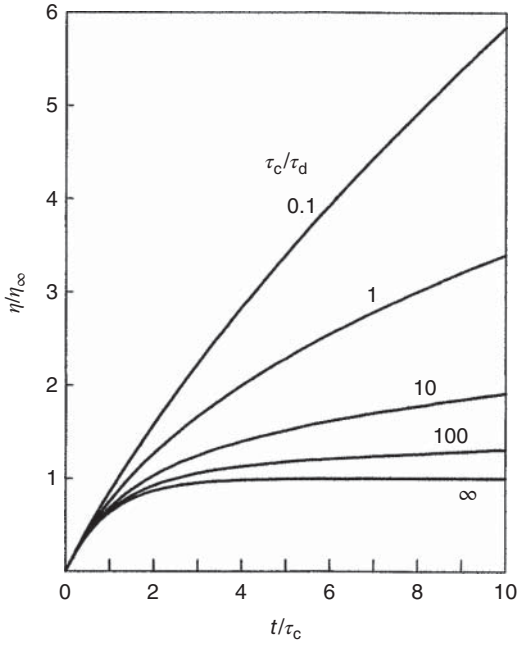


Figure 13.2 Variation of the overpotential with time during a galvanostatic transient, for different values of the parameter τ_c/τ_d. When $\tau_c/\tau_d \geq 10^3$, the reaction is under purely activation control, and η stops changing with time at about a value of $t/\tau_c \geq 3$.

and the characteristic time constant for the diffusion process, τ_d, given by

$$\tau_d^{1/2} \equiv \frac{RT}{(nF)^2}\frac{2C_{dl}}{c_b D} \tag{13.7}$$

It may look unusual that the product of capacitance and resistance would have the units of time, but this can be readily proven, by substituting the appropriate units in Eq. (13.6). Thus,

$$C_{dl} \times R_F \rightarrow \frac{\mathrm{F}}{\mathrm{cm}^2} \times \Omega \ \mathrm{cm}^2 \rightarrow \frac{\mathrm{C}}{\mathrm{V}} \times \Omega \rightarrow \frac{\mathrm{A} \times \mathrm{s} \times \Omega}{\mathrm{V}} \rightarrow \mathrm{s} \tag{13.8}$$

A similar analysis shows that τ_d also has the units of time, thus the ratio of τ_c/τ_d used in Figure 13.2 is dimensionless. Since τ_c is inversely proportional to the exchange current density, while τ_d is independent of it, a large value of τ_c/τ_d indicates that the rate of charge transfer is low, and vice versa.

13.1.4.2 The Double-Pulse Galvanostatic Method

Consider the equivalent circuit shown in Figure 13.3, ignoring for the moment the Warburg impedance associated with mass-transport limitation. When a galvanostatic pulse is applied to such a circuit, the response is that shown in Figure 13.3. The equation describing the change of overpotential with time during the transient is

$$\eta = \eta_\infty[1 - \exp(-t/\tau_c)] + jR_S \tag{13.9}$$

where the term jR_S is the Ohmic overpotential (i.e. the current density multiplied by the uncompensated solution resistance between the working and the reference electrodes).

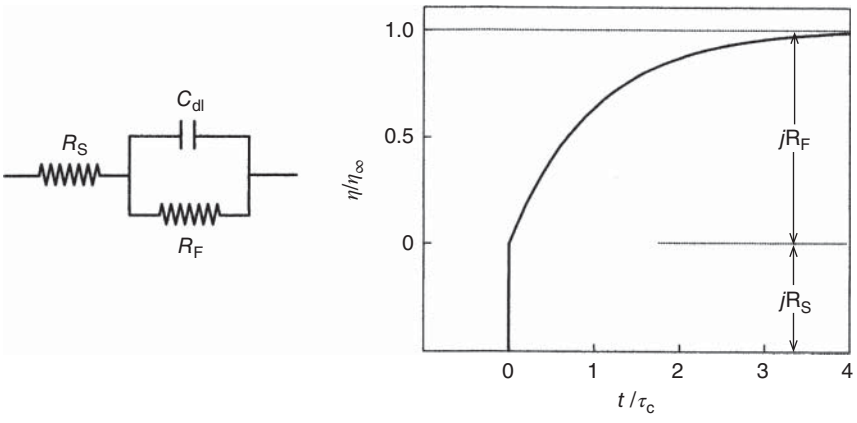


Figure 13.3 The response of the overpotential to a galvanostatic transient under purely activation controlled conditions. Presented in dimensionless coordinates, as η/η_∞ versus t/τ_c, where $\tau_c \equiv R_F C_{dl}$.

There is nothing to be gained by increasing the applied current density since, as we have seen, the relaxation time is independent of it. In fact, all transients taken at any current density are described by the single line shown in Figure 13.3, because it is plotted in dimensionless form. It is possible, however, to charge the interphase faster, by applying two consecutive pulses, the first substantially shorter but larger than the second, as shown in Figure 13.4a.

The idea behind this method is to charge the double-layer capacitor rapidly with a large current pulse j_p, and interrupt this pulse just as the overpotential has reached the correct value, corresponding to steady state at the lower current j, namely when $\eta = \eta_\infty = jR_F$ and $j \ll j_p$. One does not know this "correct" value of the overpotential, of course, since it is the quantity being measured. This disadvantage can be overcome by trial and error, as shown in Figure 13.4b. An overshoot or an undershoot can be detected, and the best value of the ratio j_p/j can be found.

The reader should be warned here that performing this type of experiment is not as easy as might be assumed, upon viewing the simulated transients shown in Figure 13.4. Factors such as electronic noise, poor impedance matching between the instruments and the electrochemical cell, and saturation of the input amplifier as a result of a large jR_S potential drop, may distort the pulse and make the determination more difficult, and sometimes impossible. Chemical factors, such as modification of the surface during the initial pulse, may also make it hard to choose the correct value of j_p/j. Most of these problems can, however, be overcome by following correct experimental procedures, and the double pulse method can yield very useful kinetic data for fast electrode reactions.

13.1.4.3 The Coulostatic (Charge-Injection) Method

Consider the application of a very short current pulse to the interphase. The charge $q_p = j \times t_p$ injected during the pulse changes the potential across the

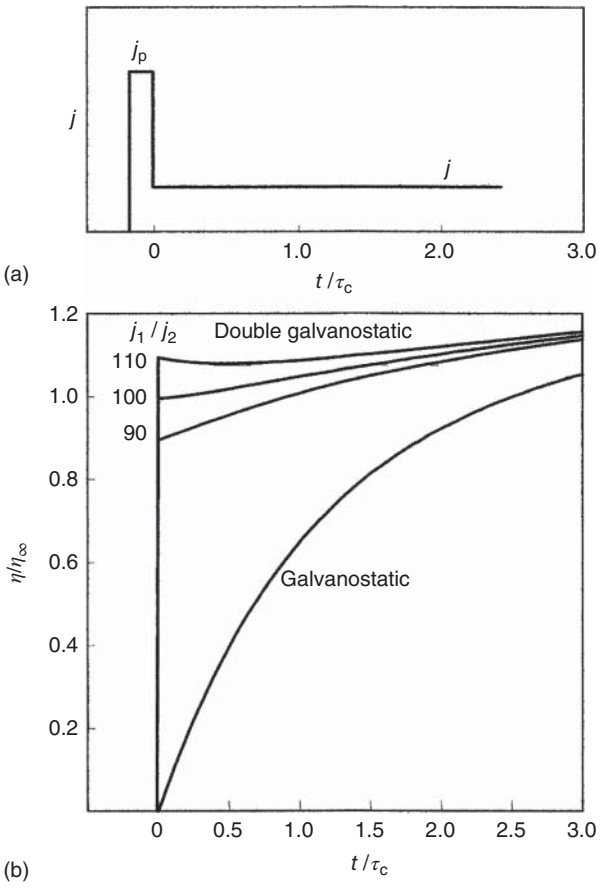


Figure 13.4 Simulated representation of the double-pulse galvanostatic method. (a) the pulse shape and (b) the response , calculated for τ_c/τ_d. A single-pulse galvanostatic transient for the same value of τ_c/τ_d is also shown for comparison. The optimal ratio of currents in this particular case turns out to be in the range of $j_p/j = 100 - 110$.

double-layer capacitance by an amount of $\Delta E_{\text{pulse}} = q_p/C_{dl}$. Starting from the open-circuit potential, this will be equal to the overpotential η_0, as seen in Figure 13.5. We use here the subscript *zero*, because this is the initial overpotential (at $t = 0$) for the open-circuit decay transient studied in a coulostatic experiment.

We shall treat here the simplest case, in which the pulse duration is very short compared to the time constant for charging the double-layer capacitance, ($t_p/\tau_c \ll 1$), and diffusion limitation can be ignored. Under such conditions, no faradaic reaction takes place during the charging pulse. Once on open circuit, the capacitor will be discharged through the faradaic resistor, R_F. It is easy to derive the form of the decay transient. On the one hand, the current is given by

$$j = -C_{dl}(d\eta/dt) \tag{13.10}$$

On the other hand, since we are dealing with small perturbations, it is also given by

$$\frac{j}{j_0} = \frac{j_0 nF\eta}{\nu RT} = \frac{\eta}{R_F} \tag{13.11}$$

Remember that the current density in Eq. (13.11) is an *internal current*, since the decay of overpotential is followed at open circuit. It is also interesting to note here that during the charging pulse, C_{dl} and R_F are effectively connected in parallel, while during open-circuit decay the same two circuit elements must be considered to be connected in series, and the same (internal) current is flowing through both.

Combining Eqs. (13.10) and (13.11) and rearranging, we can write

$$\int_{\eta_0}^{\eta} d\ln\eta = -\frac{1}{\tau_c}\int_0^t dt \tag{13.12}$$

which shows that the overpotential decays exponentially with time, following the equation:

$$\frac{\eta}{\eta_0} = \exp\left(-\frac{t}{\tau_c}\right) \tag{13.13}$$

The relaxation time for charge transfer, τ_c, can be obtained from the slope of a plot of ln η versus t. The intercept at $t = 0$ yields $\eta_0 = q_p/C_{dl}$, from which C_{dl} is obtained, since q_p is the experimentally controlled parameter. With C_{dl} known, R_F and j_0 can be obtained from τ_c.

If diffusion limitation is considered, the overpotential decays more slowly, as shown in Figure 13.5. This should be evident, since the Warburg impedance shown in Figure 13.1 is added in series with the faradaic resistance, R_F. In this case, the plot of ln η versus t is not linear, and a more complex mathematical treatment, taking into account the diffusion equations, must be applied to calculate the kinetic parameters.

The unique feature of the coulostatic method is that measurements are made at open circuit. This leads to two important consequences:

(1) Since the charge is injected in a very short time (preferably, 1 μs or less), measurement can often be completed before diffusion limitation has become significant. In this respect, the charge-injection (coulostatic) method is similar to the double-pulse galvanostatic method, except that one has more freedom in the choice of the parameters of the pulse, since there is no need to match it to the second pulse.

(2) Electrode reactions can be studied in poorly conducting solutions, since there is no error due to the jR_S potential drop (also referred to as the *resistance overpotential*) in the course of an open-circuit measurement. This feature may be particularly useful for studies in nonaqueous solutions and at low temperatures. Although this is fundamentally correct, there are practical limitations to its applicability. To show an extreme example, one cannot follow the open-circuit decay of potential over a range of 10 mV, if the jR_S potential change during the pulse is, say, 10 V. Thus, even though

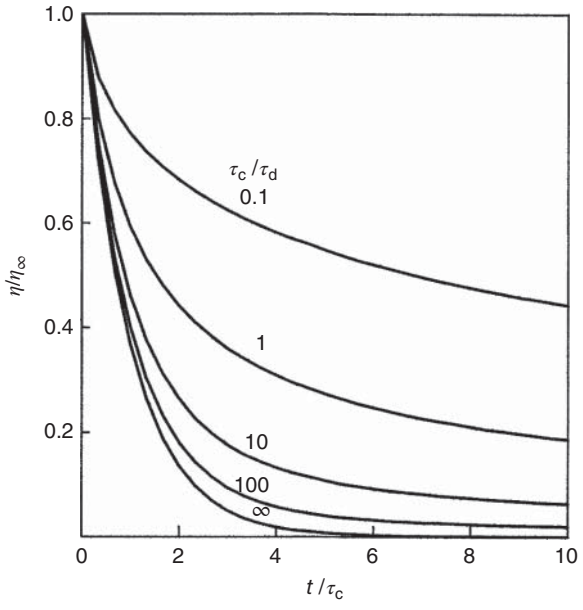


Figure 13.5 The coulostatic method (charge injection is followed by open-circuit decay). At low values of τ_c/τ_d, the decay transient is slowed down by partial diffusion control, as expected from the equivalent circuit shown in Figure 13.1.

measurements are taken at open circuit, it is clearly advantageous to use conducting solutions whenever possible, and this becomes essential if we wish to study the rate of fast reactions. These considerations become even more critical when the effects of diffusion limitation are included.

13.2 The Diffusion Equation

13.2.1 The Boundary Conditions of the Diffusion Equation

In the study of the diffusion of species to and from the electrode surface, we use the notation $c_i(t,x)$ to describe the concentration of the ith species as a function of time and distance from the electrode surface. The time-dependent diffusion equation, in its general form, is written as

$$[\partial c(x,t)/\partial t] = \nabla^2[Dc(x,t)] \tag{13.14}$$

where ∇^2 is the *Laplace operator*, corresponding to the second derivative of the concentration with respect to distance, in the appropriate coordinates. Two assumptions are commonly made in electrochemistry in order to simplify Eq. (13.14): (i) The diffusion coefficient, D, is assumed to be constant, independent of concentration, and (ii) The equation is solved in one dimension (perpendicular to the surface), assuming the so-called *semi-infinite linear diffusion*. With these assumptions, Eq. (13.14) is simplified to

$$[\partial c(x,t)/\partial t] = D[\partial^2 c(x,t)/\partial x^2] \tag{13.15}$$

How severe are these assumptions? The concentration of the electroactive species in the Nernst diffusion layer can vary from zero (at $x=0$ and $j=j_L$) to

the bulk concentration (for $x = 0$ and $j = 0$). The bulk concentration is typically a few millimole. Since measurements are conducted in the presence of a large excess of supporting electrolyte, this represents a very small change in the *total* concentration, and the variation in D as a function of the distance from the surface can be considered to be negligible.

What is *semi-infinite* in this case? How far is *infinity*, in the context of diffusion of a species in an aqueous solution? It should be far enough from the surface for the concentration to have reached its value in the bulk. This applies when x in Eq. (13.15) is in the range of 5δ. Employing the equation $\delta = \sqrt{\pi Dt}$, we find that $\delta \approx 6 \times 10^{-2}$ cm at 100 seconds, hence, *infinity* lies no more than 0.3 cm away from the surface. One does not even have to use a planar electrode in order to achieve one-dimensional (planar) diffusion. A cylindrical electrode (i.e. a wire) or a spherical electrode will also look "planar" as long as the Nernst diffusion layer thickness is small compared to the radius of curvature of the electrode. On the other hand, a micro-electrode, typically having a radius of $r \leq 10\,\mu$m, will not follow the equations for semi-infinite linear diffusion, as discussed in Section 13.3.

If the conditions for reaching an "infinite" distance from the surface are so easy to implement, why is it called "*semi-infinite*"? Because we deal with the direction perpendicular to surface into the solution, not into the electrode, namely, only *half* of the space is taken into consideration.

There are two aspects of solving the diffusion equation. One must first set the initial and boundary conditions, then find a mathematical procedure for solving the equations. Here, we shall concentrate on the former aspect. Mathematical techniques for solving the diffusion equation are discussed in many textbooks, since this problem is not unique to electrochemistry.

The initial and boundary conditions under which the diffusion equation is solved define, in mathematical language, the kind of experiment being performed and the initial state of the system being studied. It is important to realize that the resulting equations hold true only if these initial and boundary conditions have been maintained. This can be explained with the use of a few examples.

The equation to be solved, Eq. (13.15), is a second-order differential equation in two variables. It requires, therefore, three initial and boundary conditions. Consider a simple reaction of the type

$$\mathrm{Ox} + n\mathrm{e}_{\mathrm{M}}^{-} \rightleftharpoons \mathrm{Red} \tag{13.16}$$

in which both the oxidized and the reduced form are in solution. Actually, one must solve two similar diffusion equations simultaneously, one for the reactants and one for the products. Thus, there are six initial and boundary conditions that must be defined, which are discussed further.

13.2.1.1 Potential Step, Reversible Case (Chrono-Amperometry)

The initial conditions for this and all other cases to be discussed below are as follows:

$$c_{\mathrm{Ox}}(x, 0) = c_{\mathrm{b,Ox}} \quad \text{and} \quad c_{\mathrm{Red}}(x, 0) = c_{\mathrm{b,Red}} \tag{13.17}$$

In words, these equations state that the concentrations of both reactant and product are uniform everywhere in solution at $t = 0$ (i.e. before we have applied

the pulse). For simplicity, we shall assume here that there is no product initially in solution, but this is not essential. These initial conditions may seem self-evident, but they really are not. In particular, when the experiment is conducted by applying a series of pulses (e.g. each at a different potential), care must be taken to let the solution reach homogeneity before each pulse – for instance, by stirring for a short time and then allowing the solution to become completely quiescent.

The next two equations arise from the condition of *infinity*. They are written as follows:

$$c_{\mathrm{Ox}}(\infty, t) = c_{\mathrm{b,Ox}} \text{ and } c_{\mathrm{Red}}(\infty, t) = c_{\mathrm{b,Red}} = 0 \tag{13.18}$$

Since, as we have shown, "infinity" is less than 0.3 cm away during diffusion-controlled mass-transport processes, these conditions generally apply, except when a conscious effort is made to place the working and counter electrodes very close to each other, as in thin-layer cells.

Next, we have an equation of mass balance, written as follows:

$$D_{\mathrm{Ox}}[\partial c_{\mathrm{Ox}}(0, t)/\partial x] + D_{\mathrm{Red}}[\partial c_{\mathrm{Red}}(0, t)/\partial x] = 0 \tag{13.19}$$

For the reaction assumed here, one molecule of the reduced form (Red) is produced for each molecule of the oxidized form (Ox), which has been consumed; hence, the flux of the oxidized species reaching the surface must be equal to the flux of the reduced species leaving it. Since the two fluxes are in opposite directions, their sum must be zero. This boundary condition leads to some restrictions on the use of the resulting diffusion equation. Equation (13.19) is valid only if both species are soluble. It does not apply, for example, to an electroplating process, because the product stays on the surface. Even for a reaction such as bromine evolution, where both reactants and products are soluble, Eq. (13.19) would have to be slightly modified, to take into account the fact that two reacting species combine to form a single molecule of the product. The diffusion equation can, of course, be solved for these and other cases as well. The point we want to emphasize is that, before applying a diffusion equation found in the literature, it is important to know the boundary conditions under which that particular equation was solved, to ensure that it applies to the experiment being analyzed.

The sixth and last boundary condition follows from the assumption of reversibility. If the reaction is very fast, the reactants and products *at the electrode surface* are in equilibrium at all times, and their concentrations will conform to the Nernst equation. This boundary condition can be written as

$$E = E^0 + (RT/nF) \ln \left[\frac{c_{\mathrm{Ox}}(0, t)}{c_{\mathrm{Red}}(0, t)} \right] \tag{13.20}$$

Often, this equation is written in a different form as

$$\Theta = \frac{c_{\mathrm{Ox}}(0, t)}{c_{\mathrm{Red}}(0, t)} = \exp \left[\frac{nF}{RT} (E - E^0) \right] \tag{13.21}$$

Having established the physical conditions and the six initial and boundary conditions, one can proceed to solve the diffusion equation. We shall skip the mathematical derivation and proceed directly to the solution, applicable to a potential step under reversible conditions, which is

$$j = \left[\frac{nFDc_0}{\sqrt{\pi Dt}}\right] x\left(\frac{1}{1+\Theta}\right) \tag{13.22}$$

As the diffusion-limited current density is reached, the concentration of reactant at the surface is reduced to zero, therefore $\Theta = 0$. Substituting in Eq. (13.22) we have

$$j_d = \frac{nFDc_b}{\sqrt{\pi Dt}} \tag{13.23}$$

Combining Eqs. (13.22) and (13.23), we can relate the current density to potential and time by the simple equation

$$j(t) = \frac{1}{1+\Theta} j_d(t) \tag{13.24}$$

Solving the diffusion equation under the foregoing initial and boundary conditions yields the concentration profile near the electrode surface, as a function of time:

$$\frac{c(x,t)}{c_b} = \text{erf}\left[\frac{x}{\sqrt{4Dt}}\right] \tag{13.25}$$

Plots of the dimensionless concentration c/c_b as a function of distance x at different times are shown in Figure 13.6. The gradual development of the Nernst diffusion layer with time can be clearly seen. The concentration profile very close

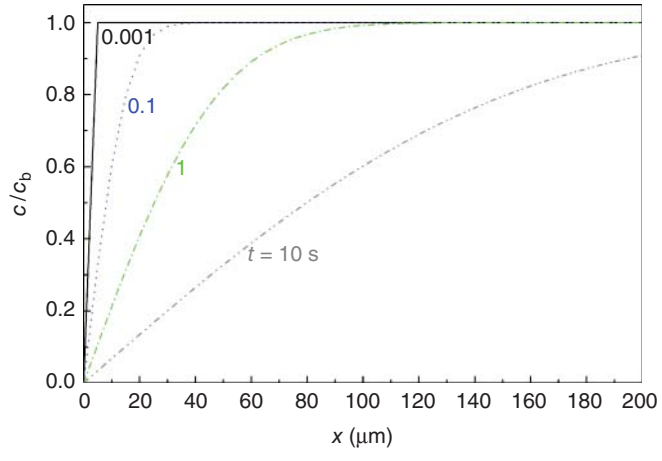


Figure 13.6 Evolution of the concentration profile with time near the electrode surface, just after the potential has been stepped to the limiting current region. $D = 7 \times 10^{-6}$ cm^2 s^{-1}

to the surface is linear, but a deviation from linearity is observed farther away, as the concentration approaches its bulk value.

13.2.1.2 Potential Step, High Overpotential Region (Chrono-Amperometry)

Our second example is also a potential step experiment. Here, however, it is assumed that the kinetics of the reaction is slow, and equilibrium is not maintained at the interphase. Moreover, we assume that the potential range studied is far from equilibrium, so that only the forward reaction needs to be considered. In the original literature this condition was referred to as the *totally irreversible case*, a term which we consider to be rather misleading, because the course of the reaction can be reversed, if a potential step in the opposite direction is applied.

The first five initial and boundary conditions of the differential equation listed above remain unchanged, only the sixth, Eq. (13.20), is different. The new boundary condition is obtained by relating the current density to the heterogeneous rate constant of the forward reaction, k_h, and to the flux of reactant at the electrode surface:

$$\frac{j}{nF} = k_h c(0,t) = D\left(\frac{\partial c(0,t)}{\partial x}\right) \tag{13.26}$$

The solution of this diffusion equation yields the current density as a function of time and potential

$$j(t) = nFk_h c_b \exp(\lambda^2)\mathrm{erfc}(\lambda) \tag{13.27}$$

in which the dimensionless parameter λ is given by

$$\lambda \equiv k_h\sqrt{t/D} \tag{13.28}$$

The heterogeneous rate constant, k_h, depends on potential exponentially, following a relationship similar to the Tafel line, namely

$$k_h = k_h^0\left[-\frac{\alpha F(E-E^0)}{RT}\right] \tag{13.29}$$

Combining Eq. (13.27) with the expression for the diffusion-limited current density, which we have obtained earlier (see Eq. (13.23)), and with Eq. (13.28), we have

$$\frac{j}{j_d} = F_1(\lambda) = \pi^{1/2}\lambda\,\exp(\lambda^2)\mathrm{erfc}(\lambda) \tag{13.30}$$

The function $F_1(\lambda)$, which has been tabulated in detail in the literature, can be used to obtain λ for any given ratio of j/j_d. In this way, k_h, which is proportional to the *activation-controlled* current density, can be evaluated as a function of the potential. A plot of ln k_h (or ln λ) versus E is equivalent to the traditional Tafel plot, in which ln j_{ac} is plotted versus E or η.

What are the limitations imposed on the validity of Eqs. (13.28) and (13.29)? In writing the sixth boundary condition we made the assumption that the reaction is first-order with respect to the reactant. This is a serious limitation, because in a study of the mechanism of electrode reactions, the reaction order is one of

the quantities we wish to determine experimentally. Obviously, the values of k_h obtained from Eq. (13.28) in electrolytes containing different concentrations of the reactant cannot be used to evaluate the reaction order, since this equation is valid only if the reaction order is unity.

13.2.1.3 Current Step (Chronopotentiometry)

Our third example, in which a current step is applied, could be called a galvanostatic experiment, in the sense that the current, rather than the potential, is the externally controlled parameter.

The first five initial and boundary conditions of the diffusion equation remain unaltered, and it is again the sixth that must be changed, to make the result applicable to this particular experimental technique. Since the current is externally controlled, one controls in effect the flux at the electrode surface. This is expressed mathematically by the equation

$$\frac{j}{nF} = D\left[\frac{\partial c(0,t)}{\partial x}\right] \tag{13.31}$$

Note that the heterogeneous rate constant, k_h, is not part of this equation, because the reaction is *forced* to proceed at a rate determined by the applied current.

Solving the diffusion equation, one obtains the concentration as a function of time:

$$\frac{c}{c_b} = 1 - (t/\tau)^{1/2}[\exp(-\chi^2) - \pi^{1/2}\chi\,\mathrm{erfc}(\chi)] \tag{13.32}$$

where the dimensionless distance parameter, χ, is given by

$$\chi = x/(4Dt)^{1/2} \tag{13.33}$$

and τ is the *transition time*, defined by

$$\tau^{1/2} \equiv \frac{nF\sqrt{\pi D}}{2j}c_b \tag{13.34}$$

This is the *Sand equation*, derived in 1901. The transition time τ is the time taken for the surface concentration of the reacting species to be reduced to zero.

Concentration profiles calculated from Eq. (13.27) for different times are shown in Figure 13.7. The important thing to note in this figure is that in a galvanostatic experiment, the flux at the surface is constant,[1] but the concentration decreases with time. In contrast, in a potentiostatic experiment, the surface concentration is held constant, but the gradient of concentration at the surface, $-D(\partial c/\partial x)_{x=0}$, decreases with time, as seen in Figure 13.6. A chrono-potentiometric experiment must be performed in quiescent solutions. Thus, while the concentration at the surface declines to zero, the bulk concentration is unchanged, as prescribed by the boundary conditions, Eq. (13.18). The meaning of the transition time τ may be clarified by considering Figure 13.8, which displays a transient obtained

1 The flux is proportional to the gradient of concentration at the electrode surface, namely to $(\partial c/\partial x)_{x=0}$, which is constant in this type of experiment.

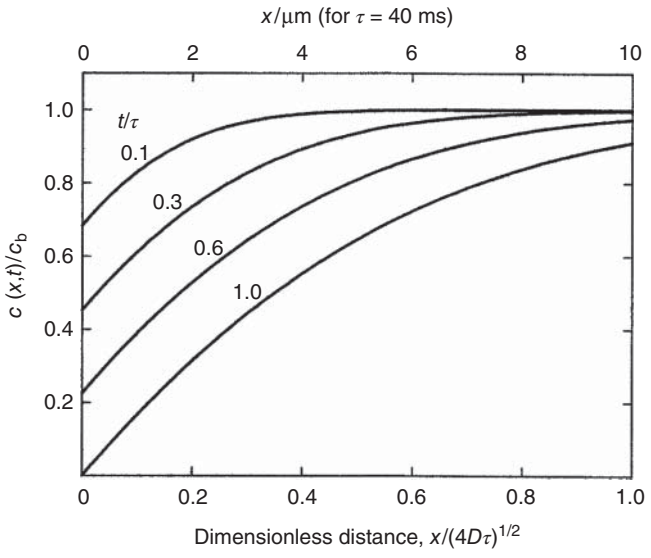


Figure 13.7 Development of the concentration profile with time during a constant-current (chronopotentiometric) transient. $\tau = 40$ ms, $D = 6 \times 10^{-6}$ cm^2 s^{-1}. The distance x is given in dimensionless form at the bottom and in micrometers at the top.

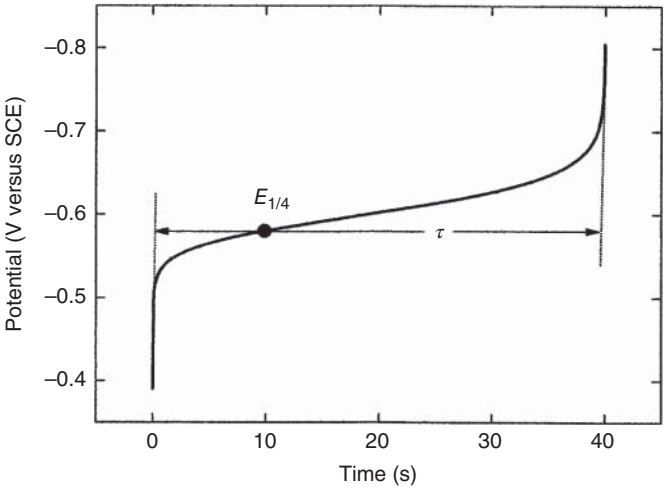


Figure 13.8 Chrono-potentiometric transient (j = const.) for a reversible reaction, showing the meaning of the transition time τ and of $E_{1/4}$ ($\tau = 40$ ms, $j = 0.16$ mA cm^{-2}, $c_b = 2$ mM Cd^{2+} in 1 M KNO_3).

for a reversible reaction. The corresponding equation describing the variation of potential with time is

$$E = E_{1/4} + \ln\left(\frac{\tau_{1/2}}{t_{1/2}} - 1\right) \tag{13.35}$$

For a slow reaction, the shape of the transient is different, but the transition time τ is independent of the kinetics of the reaction, because application of a constant current forces the reaction to proceed at the same rate, irrespective of its heterogeneous rate constant.

The quarter-wave potential $E_{1/4}$ used in Eq. (13.35) is equal to the polarographic half-wave potential and is therefore characteristic of the nature of the electroactive species in solution.

13.3 Microelectrodes

13.3.1 The Unique Features of Microelectrodes

So far, we have restricted our discussion of diffusion-controlled processes to the case of *semi-infinite linear diffusion*, which corresponds to a planar electrode in a cell where the solution extends to infinity. It has already been pointed out that the word "infinity" should not alarm us, since the dimension should be "infinitely large" only compared to the Nernst diffusion-layer thickness, δ, which is typically less than 0.3 cm. There may be always edge effects at the periphery of the electrode, but their influence on the observed current–potential relationship is usually negligible as long as the characteristic length of the electrode (i.e. the radius of a circle or the shorter side of a rectangle) is large compared to δ. As one decreases the size of the electrode, edge effects become more pronounced. Eventually, when a microelectrode is considered, edge effects become predominant. A microelectrode can be considered to be "all edge." This depends, of course, on the actual size of the microelectrode and on the timescale used. For $r = 10$ mm, the Nernst diffusion layer thickness will be equal to the radius after about 30 ms. For an ultramicroelectrode (UME) of $r = 0.25$ μm, the same is true in less than 20 μs.

The situation at a miniature disc microelectrode embedded in a flat insulator surface (such as an RDE of very small diameter) can be approximated by spherical symmetry, obtained for a small sphere situated at the center of a much larger (infinitely large, in the present context) spherical counter electrode. How will the change of geometry influence the diffusion-limited current density? This is shown qualitatively in Figure 13.9.

For a planar electrode, the diffusion-limited current density decreases with $t^{-1/2}$. In the spherical configuration, the electroactive material diffuses to each segment on the surface from a cone of given solid angle. Thus, while δ increases with time, the cross-section for diffusion increases with δ^2. As a consequence, the bottleneck for the rate of diffusion is very close to the surface (at approximately $\delta/r \leq 5$, leading to a diffusion-limited current density that becomes independent of time, as we shall see further.

The effect of geometry on the resistivity can be understood in a similar way. The real bottleneck is very close to the surface, where the resistance is high and the cross-section for conductance (which is the equivalent of the flux in the case of diffusion), is small. Farther out, the cross-section increases with the square of the distance. Hence, the contribution of this region to the total resistance soon becomes negligible, as seen in Figures 3.3 and 3.4.

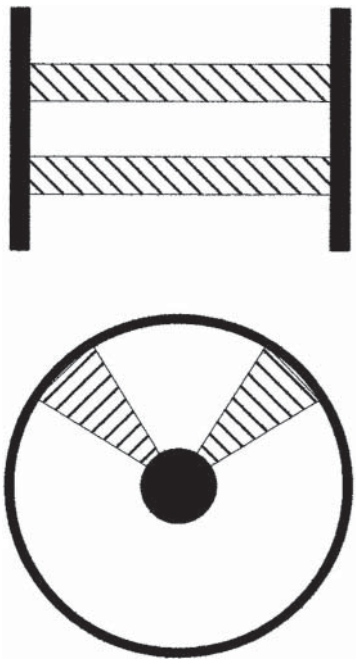


Figure 13.9 Schematic representation of planar and spherical geometry for diffusion (the latter is shown in two dimensions, for simplicity). Hatched areas represent lamina and cones from which electroactive material can diffuse to the surface.

13.3.2 Enhancement of Diffusion at a Microelectrode

The response of a spherical electrode to a potential-step function in the limiting current region is given by

$$j_d = nFDc_b \left[\frac{1}{(\pi Dt)^{1/2}} + \frac{1}{r} \right] \tag{13.36}$$

At long times, spherical diffusion is predominant, the current becomes independent of time, and Eq. (13.36) takes the form

$$j_d = \frac{nFDc_b}{r} \tag{13.37}$$

Thus, in a spherical field of diffusion (which is achieved for a microelectrode after a time determined by its radius), one obtains an equation similar to that given for semi-infinite linear diffusion, except that the radius of the electrode plays the role of the Nernst diffusion-layer thickness, and the limiting current density is independent of time. The validity of Eq. (13.37) is one of the incentives for fabricating UMEs. Thus, for example, the limiting current obtained on an RDE operated at 10^4 rpm can be equaled by the current at a microelectrode of $r = 5$ μm in a quiescent solution. Such a device is relatively easy to fabricate and would not be considered to be an *ultra*-microelectrode. Smaller electrodes, having a radius of 0.25 μm (or even less) have been prepared in several laboratories. Using Eq. (13.37) we note that the limiting current density at such an electrode is about 0.8 A cm^{-2} when $n = 2$ and $c_b = 10$ mM. Such limiting current densities, which cannot be reached at steady state by any other method, substantially increase

the range over which the current–potential relationship can be observed under activation-controlled conditions, as discussed in Section 13.1.1.

Another advantage of having a very large limiting current density at steady state is in the analysis of trace elements. Using Eq. (13.37) for the same sized electrode, we obtain a current density of about 8 μA cm^{-2} for a concentration of 0.01 ppm (assuming a molecular weight of 100 for the purpose of this calculation). Thus, measurements in the part per billion region should be possible with UMEs.

When a UME is used as an electroanalytical tool, the diffusion-limited current density is not affected by the rate of flow, which typically creates a diffusion layer thickness of the order of 50–100 μm. This can be a great asset for online monitoring in industrial applications, where the flow rate may fluctuate and would otherwise have to be measured and corrected for.

13.3.3 Reduction of the Solution Resistance

The jR_S potential drop due to the uncompensated solution resistance associated with different geometries was discussed in Section 3.2.3. For a spherical electrode, which is of interest here, we can write

$$E_{jR_S} = j\frac{d}{\kappa}\left[\frac{r}{r+d}\right] \tag{13.38}$$

where κ is the specific conductivity and d is the distance of the probe from the electrode surface. Of interest to us here is the limiting form of this equation, for large values of d, namely, far away from the electrode. In this case, Eq. (13.36) is simplified to

$$jR_S = j\frac{r}{\kappa} \tag{13.39}$$

showing that the resistance (in units of Ω cm^2, i.e. normalized for unit surface area), is proportional to the radius of the UME.

As usual, a numerical example may help to illustrate the advantage of UMEs, from the point of view of solution resistance. In Section 13.3.2, we obtained a limiting current density of 0.8 A cm^{-2} for an electrode having a radius of 0.25 μm, in a 10 mM solution of the reactant. If we assume a specific conductivity of $\kappa = 25$ mS cm^{-1}, the solution resistance R_S according to Eq. (13.39) is 1×10^{-3} Ω cm^2. We assumed here a solution of medium specific conductivity, and yet arrived at an Ohmic potential drop of less than 1 mV at a very large current density of 0.8 A cm^{-2}. Thus, using a UME extends the range of measurable current densities because: (i) the limiting current density is inversely proportional to the radius, and (ii) the uncompensated resistivity is proportional to the radius. More precisely, we could say that both the *diffusion-limited current density* and the *conductivity* are inversely proportional to the radius.

13.3.4 The Choice between Single Microelectrodes and Large Ensembles

Microelectrodes have many advantages over regular-sized electrodes, but they also have two major disadvantages. First, since the electrode is very small, the

total current flowing in the circuit is minute, and might be difficult to measure accurately. The UME just discussed has a total surface area of about 2×10^{-9} cm^2. For a current density of 1.0 mA cm^{-2}, the total current observed is hence only 2×10^{-12} A. Although this is measurable in the research laboratory environment, the measurement is by no means easy, and accuracy might be limited. This makes the use of UMEs for routine applications, particularly in industrial environments, inconvenient. The second disadvantage entails the extremely high volume-to-surface ratio, which makes it impossible to purify solutions to a level that would ensure that impurities will not accumulate on the surface during measurement, as discussed in Section 7.1.2. Thus, for a regular electrode, one may have a volume-to-area ratio of 10 cm^3 cm^{-2}. For a thin-layer cell, this ratio could be as low as 10^{-3} cm^3 cm^{-2}, while for a UME the ratio is of the order of 10^8 cm^3 cm^{-2}. This is an inherent difficulty, which cannot be overcome by improved instrumentation. These numbers mean that, given a desired level of purity *of the surface*, the allowed level of impurities in solution would have to be 9 orders of magnitude lower for the UME than for a regular electrode. Bearing in mind that proper electrochemical measurements must be conducted in highly purified solutions, even when electrodes of macroscopic dimensions are employed, the level of purity needed for work with a single UME is clearly not achievable.

One way around this problem, which retains most of the advantages of UMEs while largely overcoming their disadvantages, is to use ensembles of UMEs. Suppose that the surface of an insulator is dotted with a regular array of conducting spots that serve as the UMEs, which are all connected at the back to a common current collector, as shown in Figure 13.10.

Assuming a uniform distribution, the fraction of the surface that is active is given by $(d/L)^2$ where d is the diameter of each electrode and L is the distance between their centers. Designing such an ensemble of microelectrodes, one must compromise between the desire to make the ratio $(d/L)^2$ as small as possible (to decrease the overlap between the diffusion fields of the individual electrodes), and the desire to make it as large as possible (to increase the total active area). Values of $(d/L)^2$ in the range of 0.10–1.1 times the active area seem to be a reasonable

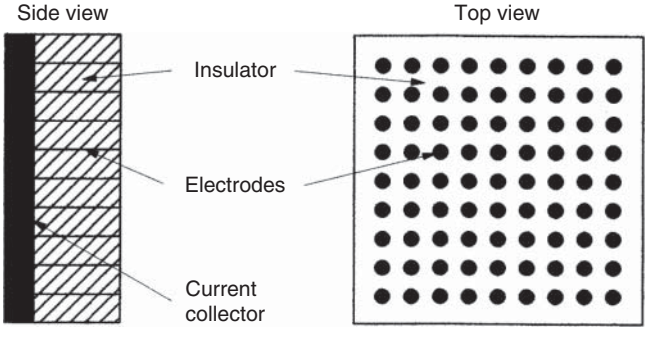


Figure 13.10 An ensemble of ultramicroelectrodes (UMEs). Note that the distance between the electrodes is large compared to their diameter.

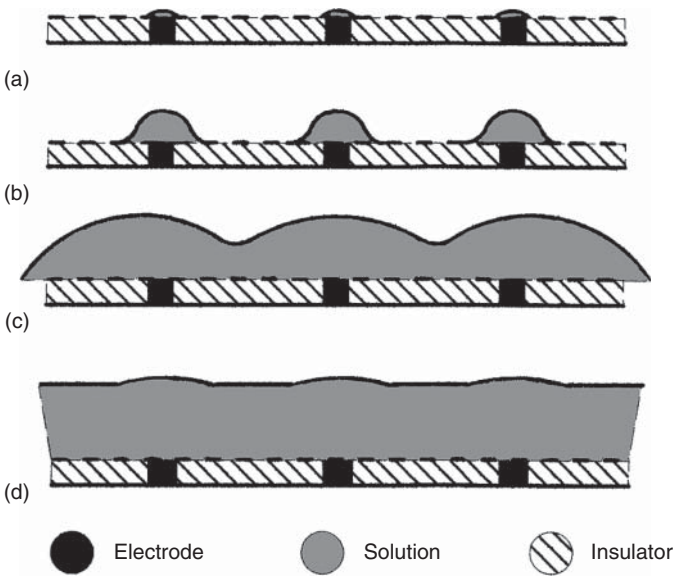


Figure 13.11 Development of the diffusion field near the surface of an ensemble of microelectrodes. (a) semi-infinite linear diffusion; (b) spherical diffusion with no overlap; (c) spherical diffusion with substantial overlap; (d) total overlap, equivalent to planar diffusion to the whole surface.

choice. Figure 13.11 shows qualitatively the way the diffusion layer would develop as a function of time in an ensemble of microelectrodes.

Further Reading

1 Bard, A.J., Inzelt, G., and Scholz, F. (ed.) (2012). *Electrochemical Dictionary*. Berlin Heidelberg: Springer-Verlag.
2 Zoski, C.G. (ed.) (2007). *Handbook of Electrochemistry*. Amsterdam, The Netherlands: Elsevier.
3 Bagotsky, V.S. (2006). *Fundamentals of Electrochemistry*, 2nd edn. New Jersey, USA: Wiley.
4 Bard, A.J. and Faulkner, L.R. (2001). *Electrochemical Methods: Fundamentals and Applications*, 2nd edn. New York, USA: Wiley.
5 Forster, R.J. (2000). Ultrafast electrochemical techniques. In: *Encyclopedia of Analytical Chemistry*, vol. 11 (ed. R.A. Meyers), 10142–10171. Chichester, UK: Wiley.
6 Aoki, K. (1993). Theory of ultramicroelectrodes. *Electroanalysis* 5: 627–639.
7 Amatore, C. (1993). Unconventional electrochemistry at ultramicroelectrodes: new approaches for investigation of chemical reactivity. In: *Molecular Electrochemistry of Inorganic, Bioinorganic and Organometallic Compounds* (ed. A.J.L. Pombeiro and J.A. McCleverty), 625–644. New York, USA: Kluwer Academic Publishers.

8 Heinze, J. (1993). Ultramicroelectrodes in electrochemistry. *Angew. Chem. Int. Ed. Engl.* 32: 1268–1288.
9 Scharifker, B.R. (1992). Microelectrode techniques in electrochemistry, Chapter 5. In: *Modern Aspects of Electrochemistry*, vol. 22 (ed. J.O.'M. Bockris, B.E. Conway and R.E. White), 467–520. New York, USA: Springer Science+Business Media.
10 Gileadi, E., Kirowa-Eisner, E., and Penciner, J. (1975). *Interfacial Electrochemistry: An Experimental Approach*. Addison-Wesley.

14

Linear Potential Sweep and Cyclic Voltammetry

14.1 Three Types of Linear Potential Sweep

Linear potential sweep is a potentiostatic technique, in the sense that the potential is the externally controlled parameter. The potential is changed at a constant rate of

$$\nu = \frac{\mathrm{d}E}{\mathrm{d}t} \tag{14.1}$$

and the resulting current is followed as a function of time.

In most cases, the potential is swept forward and backward between two fixed values, a technique referred to as *cyclic voltammetry* (CV). In this way, the current measured at a particular potential on the anodic sweep (going from negative to positive potentials) can readily be compared with that measured at the same potential on the cathodic sweep (going from positive to negative potentials). Note that the position of the peak potentials is shifted to some extent. A typical cyclic voltammogram is shown in Figure 14.1. Linear potential sweep measurements are generally of three types (see Sections 14.1.1–14.1.3).

14.1.1 Very Slow Sweeps

In the first type, the sweep rate is very low, in the range of $\nu = 0.1\text{–}5\,\mathrm{mV\,s^{-1}}$, and measurement is conducted under quasi-steady-state conditions. The sweep rate plays no role in this case, except that it must be slow enough to ensure that the reaction is effectively at steady state at each potential in the course of the sweep. This type of measurement is widely used in corrosion and passivation studies, as we shall see, and also in the study of some fuel cell reactions in stirred solutions. Reversing the direction of the sweep should have no effect on the current–potential relationship, if the sweep is slow enough. Deviations occur sometimes as a result of slow formation and/or reduction of surface oxides or passive layers. Because the sweep rate is slow, the potential is often swept only in one direction, and the experiment is then referred to as *linear sweep voltammetry* (LSV).

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

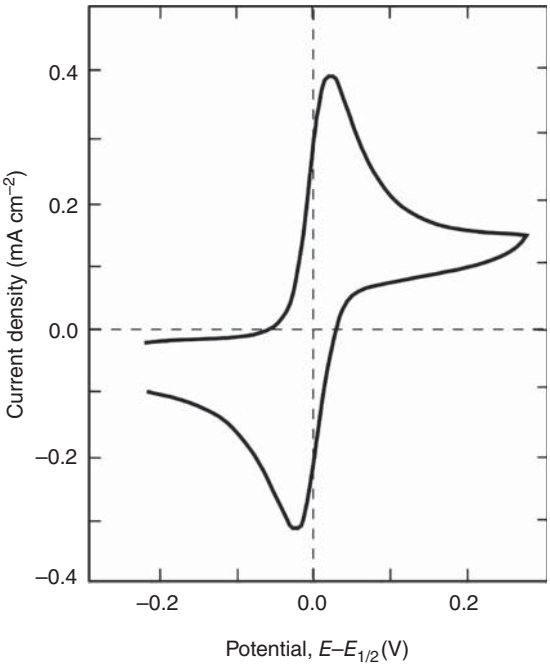


Figure 14.1 A simple cyclic voltammogram ($\nu = 0.01$ V s^{-1}, 3 mM $FeSO_4$, 1.0 M H_2SO_4).

14.1.2 Studies of Oxidation or Reduction of Species in the Bulk of the Solution

In the second type, the sweep rate is usually in the range of 0.01–10 V s^{-1}. The lower limit is determined by the need to maintain the total time of the experiment below 10–50 seconds (before mass transport by natural convection becomes significant). The upper limit is determined by the double-layer charging current and by the uncompensated solution resistance, as discussed in Section 14.2.

14.1.3 Studies of Oxidation or Reduction of Species Adsorbed on the Surface

In the third type, the redox behavior of species that are adsorbed on the surface is usually activation controlled and influenced by the remaining number of free sites on the surface. Hence, this may be conducted in stirred solutions. The typical sweep rates are also in the range of 0.01–10 V s^{-1}, but here the lower limit is determined by background currents from residual impurities in solution (and, perhaps, by the desire of the experimenter to collect more data in a given time), while the upper limit is determined by the uncompensated solution resistance and by instrumentation. The faradaic current associated with this process is proportional to the sweep rate, and so is the double-layer charging current, so that the relative effect of j_{dl} on the measured current is independent of sweep rate.

Although the amount of material adsorbed in a monolayer is very small (on the order of 1–2 nmol cm^{-2}), mass transport can be a limiting factor under certain

circumstances, when the bulk concentration of the adsorbate is low (typically, below a few millimole and/or when the sweep rate is high).

In this chapter we shall discuss only the second and third cases, since very slow sweeps are just a convenient method for scanning the potential automatically, under conditions in which the sweep rate or its direction have practically no effect on the current observed.

14.1.4 Double-Layer Charging Currents

We cannot discuss the CV and LSV methods properly without appreciating the importance of the double-layer charging current in this type of experiment. We use the simple equation

$$j_{dl} = C_{dl}(dE/dt) = C_{dl}\nu \tag{14.2}$$

If double-layer charging is the only process taking place in a given potential region, the only current observed will be j_{dl}. Such measurements can be used to determine the value of C_{dl}. In addition to the determination of C_{dl}, it is also important to determine the numerical value of j_{dl}, since this represents an error term that must be corrected for in the analysis of the current–potential relation during CV. For a very slow sweep experiment using, for example, $\nu = 1\,\text{mV s}^{-1}$ and $C_{dl} = 20\,\mu\text{F cm}^{-2}$, a value of $j_{dl} = 20\,\text{nA cm}^{-2}$ is obtained. Such value is usually negligible, even with respect to the small currents observed on passivated electrodes, and can hence be ignored. However, cyclic voltammetry is often conducted at higher sweep rates, where j_{dl} can no longer be ignored.

The interplay between double-layer charging and oxidation or reduction of an electroactive material in the bulk of the solution is illustrated in Figure 14.2. The faradaic currents shown by the solid lines are the peak currents j_p calculated according to Eq. (14.8), which is discussed in Section 14.2.

It is important to note that the faradaic current changes with $\nu^{1/2}$, while j_{dl} is proportional to ν. As a result, double-layer charging becomes more important with increasing sweep rate. The hatched area in Figure 14.2 represents the region

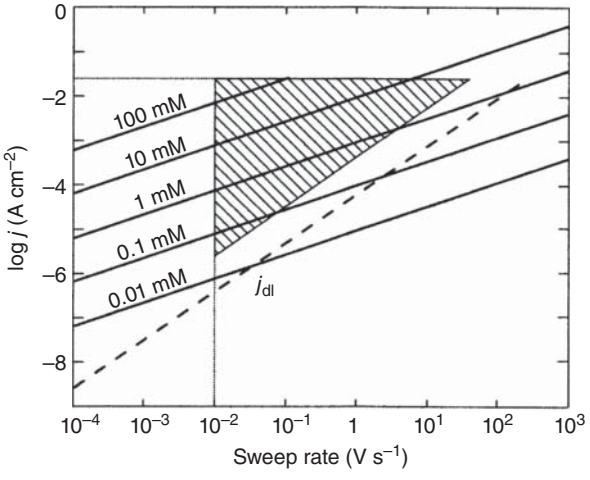


Figure 14.2 The optimum range of concentrations and sweep rates (hatched area) for measurements in cyclic voltammetry on a smooth electrode, $n = 1$. The double-layer charging current j_{dl} was calculated for $C_{dl} = 20\,\mu\text{F cm}^{-2}$.

of sweep rates and concentrations in which accurate measurements can be made, on the basis of the following assumptions:

(a) The peak current of the faradaic process being studied should be at least 10 times larger than the double-layer charging current, in order to allow reliable correction for the latter.
(b) The sweep rate should be $\nu \geq 10\,\mathrm{mV\,s^{-1}}$, chosen so as to limit errors due to natural convective mass transport.
(c) The peak faradaic current should be $j_\mathrm{p} \leq 20\,\mathrm{mA\,cm^{-2}}$, in order to limit the error due to uncompensated solution resistance.
(d) The bulk concentration should be $c_\mathrm{b} \leq 100\,\mathrm{mM}$, to ensure that it will always be possible to have an excess of supporting electrolyte, to suppress mass transport by migration.

The numerical values that have been used to construct Figure 14.2 are somewhat arbitrary. They do represent, however, the right order of magnitude, and are probably correct within a factor of 2 or 3. The important thing to learn from Figure 14.2 is the type of factors that must be considered in setting up an experiment. One could readily construct a similar diagram using somewhat different assumptions, but the conclusions would not be fundamentally different. Thus, we note that it is difficult to make measurements at concentrations below 0.1 mM or at sweep rates above $100\,\mathrm{V\,s^{-1}}$. The best concentration to use is in the range of 1–10 mM. Interestingly, this conclusion does not depend strongly on the four assumptions made above, since it refers to the "middle of the field" of applicability of the method.

An interesting point to consider is the effect of surface roughness on the range of applicability of the linear potential sweep method. The double-layer charging current j_dl in Figure 14.2 is calculated for $C_\mathrm{dl} = 20\,\mu\mathrm{F\,cm^{-2}}$, which is a relatively low value for solid electrodes, implying a highly polished surface with a roughness factor of 1.5 or less. If the roughness factor is increased, for example by using platinized platinum instead of bright platinum, the charging current could increase by a factor of typically 20–100, while the diffusion current remains essentially unchanged.

Using corresponding values of j_dl in Figure 14.2, we note that the range of applicability is seriously limited to concentrations above 1 mM and to sweep rates below a few volts per second. Thus, LSV should be conducted, whenever possible, on smooth electrodes. The application of this method to the study of porous electrodes, of the type used in fuel cells and in metal–air batteries, is rather limited.

We should perhaps conclude this section by noting that the above considerations are valid for large electrodes, operating under semi-infinite linear diffusion conditions. If microelectrodes are employed, the diffusion rate is higher and the uncompensated solution resistance is significantly lower, increasing the range of measurement by an extent which depends on the radius of the microelectrodes, as seen in Section 13.3. Indeed, one can find examples in the literature where sweep rates as high as $\nu = 1 \times 10^6\,\mathrm{V\,s^{-1}}$ were applied. Although this may be valid from the points of view of mass transport and solution resistance, one should keep in mind that the double-layer charging current density may become very large at such high sweep rates.

14.1.5 The Form of the Current–Potential Relationship

We have seen a typical j/E curve obtained in a CV experiment in Figure 14.1. This particular curve was observed in a solution containing 3 mM Fe^{2+} ions in 1.0 M H_2SO_4. Thus, the anodic peak corresponds to the simple reaction

$$Fe^{2+} \rightarrow Fe^{3+} + e_M^- \tag{14.3}$$

in which both reactant and product are stable species and both are soluble. The initial concentration of Fe^{3+} is zero, but its *surface concentration* is very close to the bulk concentration of Fe^{2+}. As a result, the cathodic reduction peak is nearly equal to the anodic oxidation peak, as long as the product of the cathodic reduction is stable in solution.

Why is a peak observed in this type of measurement? In the experiment shown in Figure 14.1 the sweep is started at a potential of −0.2 V versus the standard potential for the Fe^{2+}/Fe^{3+} couple, where no faradaic reaction takes place.[1] At about −0.05 V, the anodic current starts to increase with potential. Initially, the current is activation controlled, but as the potential becomes more positive, diffusion limitation sets in. The inverse of the current observed is equal to the sum of the inverse activation and diffusion limited currents (see Eq. (1.2)). As time goes on, the activation-controlled current increases due to the increase of potential with time, but the diffusion-controlled current decreases. As a result, the observed current increases first, passes through a maximum, and then decreases. At higher sweep rates, each potential is reached at a shorter time, when the effect of diffusion limitation is less; hence, the peak current is found to increase with sweep rate.

In CV the potential is made to change linearly with time between two set values. Often, the current during the first cycle is quite different from that in the second cycle, but after 5–10 cycles the system settles down, and the current traces the same line as a function of potential, independent of the number of cycles. This is often referred to as a "steady-state voltammogram," which is an odd name to use, considering that the current keeps changing periodically with potential and time. Nevertheless, the term is used here in the sense that *the voltammogram as a whole* is independent of time. This is a very nice experiment to perform because it tends to be highly reproducible. Moreover, the voltammograms are stable over long periods of time, sometimes in the range of hours. Interpretation of the data is another matter. Using the results for *qualitative* detection of reactions taking place in a given range of potential is fine, but quantitative treatment, using the equations developed for a single linear potential sweep is wrong because: (i) the initial and boundary conditions of the diffusion equation have changed, and (ii) convective mass transport may already play an important role.

1 In fact, there is a small cathodic current flowing, which must be attributed to the reduction of some impurity, but this is of no interest here.

14.2 Solution of the Diffusion Equations

The boundary conditions for solving the diffusion equation for linear potential sweep are really the same as those written for the potential step experiment discussed in Section 13.2, because in both cases the potential is the externally controlled parameter. As before, we can distinguish between the reversible case, in which it is assumed that the concentrations at the surface are determined by the potential via the Nernst equation (see Eq. (13.16)), and the high overpotential region, where the specific rate constant is related to the surface concentration $c(0,t)$ and to the flux at the surface, as given by Eq. (13.25). There is one important difference, however. During a potential step experiment, the potential is assumed to be constant throughout the transient, while in CV it is assumed to change linearly with time, following the simple equation

$$E = E_{\text{in}} \pm \nu t \tag{12.2}$$

where E_{in} is the initial potential at $t = 0$, which is chosen in a range where no faradaic reaction takes place. Incorporating this equation into the appropriate boundary condition will then produce a result that is applicable to the CV experiment.

The change in the boundary conditions complicates the mathematics, to the point that an explicit algebraic solution for the whole j/E curve has not been found. Numerical solutions are, however, available for the different cases. The discussion here is limited to the coordinates of the peak, namely to the values of the peak current j_{p} and the position of the peak potential E_{p}, as a function of sweep rate and the kinetic parameters of the reaction involved.

14.2.1 The Reversible Region

The peak potential in the case of a reversible linear potential sweep is given by the equation

$$E_{\text{p,rev}} = E_{1/2} \pm 1.1\frac{RT}{nF} \tag{14.4}$$

where $E_{1/2}$ is the polarographic half-wave potential, which is very close to the standard potential E^0:

$$E_{1/2} = E^0 + \frac{2.3RT}{nF}\log\left(\frac{D_{\text{Red}}}{D_{\text{Ox}}}\right)^{1/2} \tag{14.5}$$

The positive sign in Eq. (14.4) is applicable to an anodic sweep, while the negative sign applies to a cathodic peak. In either case, the peak should appear at $(28/n)$ millivolt *after* $E_{1/2}$, in the direction of the sweep. The peak potential is independent of the sweep rate in the reversible case. This characteristic can, in fact, be used as a criterion for reversibility. It is also independent of concentration. However, this is correct only if one considers a simple reaction, such as represented by Eq. (14.3). Let us consider a slightly more complex stoichiometry, such as

$$2\text{Cl}^- \rightarrow \text{Cl}_2 + 2\,\text{e}_{\text{M}}^- \tag{14.6}$$

In this case, the half-wave potential and $E_{p,rev}$ depend logarithmically on concentration:

$$E_{p,rev} = E^0 + \frac{2.3RT}{nF} \log \left(\frac{D_{Red}}{D_{Ox}} \right)^{1/2} + \frac{2.3RT}{nF} \log c_{b,Cl^-} \tag{14.7}$$

The peak current density for a simple reversible linear case (see Eq. (14.5)) is given by

$$j_{p,rev} = \left[0.44nF \left(\frac{nF}{RT} \right)^{1/2} D^{1/2} \right] \; c_b v^{1/2} \tag{14.8}$$

which can be written, for room temperature, as

$$j_{p,rev} = [2.72 \times 10^5 \, n^{3/2} D^{1/2}] \; c_b v^{1/2} \tag{14.9}$$

In this equation, $j_{p,rev}$ and c_b are given in units of A cm^{-2} and mol cm^{-3}, respectively. Also, the activity coefficients γ_{Ox} and γ_{Red} in Eqs. (14.5) and (14.7) are ignored.

14.2.2 The High-Overpotential Region

In the high-overpotential region, the peak potential depends logarithmically on sweep rate, following the equation

$$E_{p,irrev} = E_{1/2} + \frac{b}{2} \left[1.04 - \log \frac{b}{D} + \log \frac{v}{k_f^2} \right] \tag{14.10}$$

From a plot of $E_{p,irrev}$ versus $\log(v)$ one can obtain the Tafel slope, b, and from the intercept at $E_{p,irrev} = E_{p,rev}$ – the specific rate constant can be calculated. Note that the value of k_f obtained in this way is that corresponding to $E = E_{rev}$ only, but since the Tafel slope is known, k_f at any other potential can readily be calculated.

The peak current density in the high-overpotential region is given by

$$j_{p,irrev} = [3.01 \times 10^5 n \alpha^{1/2} D^{1/2}] \; c_b v^{1/2} \tag{14.11}$$

In this equation, α is the transfer coefficient, which is obtained directly from the Tafel slope. The dependence of E_p on sweep rate over a wide range is shown in Figure 14.3.

The ratio of the peak current densities in the two regions is given by

$$\frac{j_{p,irrev}}{j_{p,rev}} = 1.07 \left(\frac{\alpha}{n} \right)^{1/2} \tag{14.12}$$

Since (α/n) is usually smaller than unity, the peak current in the high-overpotential region is, as a rule, smaller than that in the reversible region. The difference is not very large, however. For $\alpha = 0.5$ and $n = 1$ the above ratio is equal to 0.78, and for $\alpha = 0.5$ and $n = 2$, it is reduced to 0.55.

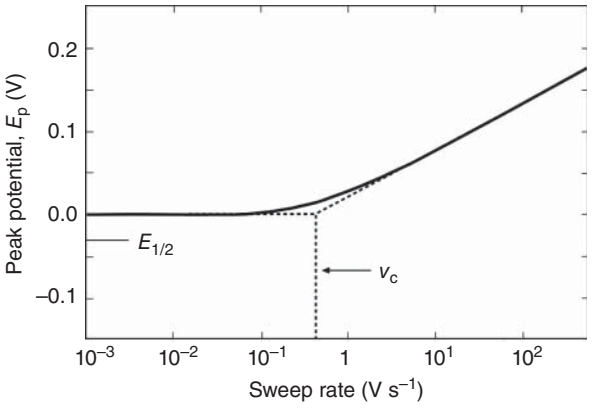


Figure 14.3 Variation of the peak potential E_p over a wide range of sweep rates, covering both the reversible and the high-overpotential regions. $E_{p,rev}$ is arbitrarily taken as zero. The "critical" sweep rate v_c is the one to be used in Eq. (14.10) to calculate k_f.

14.3 Uses and Limitations of the Linear Potential Sweep Method

Determining the peak current density in cyclic voltammetry can sometimes be problematic, particularly for the reverse sweep, or when there are several peaks that are not sufficiently separated on the scale of potential. The usual way to determine the peak currents is shown in Figure 14.4. For the forward peak, the correction for the baseline is small and does not substantially affect the result. For the two reverse peaks, however, the baseline correction is quite large and might introduce a substantial uncertainty in the value of the peak current density. In fact, there is no theory behind the linear extrapolation of the baselines shown in Figure 14.4, and this leaves room for some degree of "imaginative extrapolation." This is one of the weaknesses of cyclic voltammetry, when used as a *quantitative tool*, in the determination of rate constants and reaction mechanisms.

However, the greatest drawback in the *quantitative* use of linear potential sweep is due to the uncompensated solution resistance. We have already discussed this point in some detail with respect to galvanostatic and potentiostatic measurements. It was shown that if this uncompensated solution resistance cannot be reduced to a negligible level (either by proper cell design or by electronic means, or preferably by both), galvanostatic measurements are, as a rule, more reliable. The reason, as we have demonstrated, is that in a galvanostatic measurement the experiment is conducted correctly, even if jR_S is substantial. One is then left with the challenge of measuring a small activation-controlled overpotential on top of a large (but constant) signal due to solution resistance, which is a matter of using high-quality measuring instruments. In a potentiostatic experiment, the uncompensated solution resistance distorts the shape of the pulse, so that the experiment itself is not conducted under the presumed conditions, since the potential step is no longer a sharp step, a problem that cannot be corrected *post factum* by increasing the sensitivity and accuracy of the measuring equipment. Linear potential sweep, being a potentiostatic technique, suffers from the same drawback, but the problem in this case is even worse, as we shall see later.

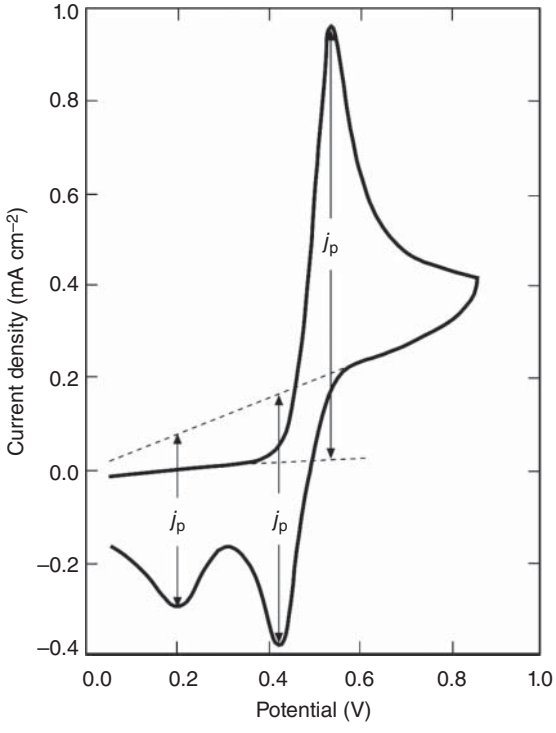


Figure 14.4 Commonly used graphical method of extrapolating the baseline, to measure the peak current densities in cyclic voltammetry.

In Figures 4.3 and 4.4 we compared the potential applied to the interphase during linear potential sweep, with and without an uncompensated solution resistance. Clearly, the error is maximal at the peak, where the current has its highest value. Just before and during the peak, the effective sweep rate imposed on the interphase is much less than that applied by the instrument. The assumption that the sweep rate is constant, which was used as one of the boundary conditions for solving the diffusion equation, does not apply in this region. In this sense, the experiment is no longer conducted "correctly."

The problem is aggravated by the common practice of extracting the kinetic information from the coordinates of the peak, namely from the values of j_p and E_p, and their dependence on sweep rate. In other words, *the information is obtained from the point at which the error is the highest*! Moreover, this is by no means a constant error. As the sweep rate is increased, the peak current increases and the error due to jR_S become worse. The effect can be quite dramatic, as seen by comparing the two cyclic voltammograms shown in Figure 4.3, which were obtained in the same solution, with and without electronic jR_S compensation. Admittedly, we have used rather extreme conditions, in which the voltammogram is visibly distorted, but it should be evident that even much smaller values of jR_S might make the results of a *quantitative* analysis questionable.

We conclude this section by noting that linear potential sweep and cyclic voltammetry are excellent *qualitative* tools in the study of electrode reactions. However, their value for obtaining *quantitative* information is rather limited.

The best advice to the novice in the field is that *cyclic voltammetry should always be the first experiment method used to study a new system, but never the last*.

14.4 Cyclic Voltammetry for Monolayer Adsorption

In Chapter 10 we discussed the concept of adsorption pseudocapacitance and its dependence on potential and the fractional coverage. The phenomenon of underpotential deposition was discussed in Chapter 11, noting that UPD is a prime example in which the adsorption pseudocapacitance plays a role. Both phenomena are studied in most cases (but not exclusively) by applying cyclic voltammetry. Here, we discuss the theory behind cyclic voltammetry associated with the above two phenomena.

14.4.1 Reversible Region

Consider a simple charge-transfer process, leading to the formation of an adsorbed intermediate, such as

$$H_3O^+ + e_M^- \rightleftharpoons H_{ads} + H_2O \tag{14.13}$$

As we sweep the potential from an initial value where $\theta = 0$ to a final value where θ is essentially unity and back to the initial value, we observe a faradaic current, associated with the formation and removal of a monolayer of adsorbed species. If we hold the potential constant at any value within this range, the current decays to zero, since the coverage is a function of the potential and does not continue to change with time at constant potential. This is the behavior characteristic of a capacitor and, in fact, the current measured during the sweep, which we denote j_ϕ, is that required to charge and discharge the adsorption pseudocapacitance C_ϕ, discussed in Section 10.2. Assuming that bulk faradaic processes and double-layer charging are negligible, we can write

$$j_\phi = q_1 \frac{\mathrm{d}\theta}{\mathrm{d}t} = q_1 \frac{\mathrm{d}\theta}{\mathrm{d}E}\frac{\mathrm{d}E}{\mathrm{d}t} = C_\phi \times \nu \tag{14.14}$$

This makes life easy, because we already know how C_ϕ depends on potential for different isotherms. If adsorption follows the Langmuir isotherm, the current during cyclic voltammetry is given by

$$j_\phi = \frac{q_1 F}{RT}\left[\frac{K_0 c_\mathrm{b} \exp(-EF/RT)}{[1 + K_0 c_\mathrm{b} \exp(-EF/RT)]^2}\right] \times \nu \tag{14.15}$$

which we obtained by substituting the value of C_ϕ from Eq. (11.25) into Eq. (14.14).

We cannot derive an explicit form of the dependence of C_ϕ on potential, but the shape of the curve can be obtained numerically from the dependence of C_ϕ on θ, on the one hand, and from the dependence of θ on E, on the other. The peak current density is obtained by combining Eq. (11.27) with (14.15):

$$j_{\mathrm{L,p}} = \left[\frac{q_1 F}{4RT}\right] \times \nu \tag{14.16}$$

for the Langmuir isotherm, and

$$j_{\mathrm{F,p}} = \left[\frac{q_1 F}{4RT}\right] \times \left[\frac{1}{1+(f/4)}\right] \times \nu \tag{14.17}$$

for the Frumkin isotherm, where $f \equiv r/RT$ is the dimensionless parameter defining the rate of change of the standard Gibbs energy of adsorption with coverage. It is important to note that the peak current density, and indeed the current density at any value of the potential, is proportional to the sweep rate ν. This makes it relatively easy to distinguish between activation-controlled surface processes and diffusion-controlled bulk processes, for which the peak current is proportional to $\nu^{1/2}$.

The dependence of the potential on the surface coverage according to Frumkin is

$$\left[\frac{\theta}{1-\theta}\right] \exp\left(\frac{r}{RT}\theta\right) = K_0 c_{\mathrm{b}} \exp\left(-\frac{F}{RT}E\right) \tag{14.18}$$

Bearing in mind that the maximum value of the pseudocapacitance always occurs at $\theta = 0.50$, irrespective of the value of the parameter f (see Figure 14.5), this yields

$$E_{\mathrm{p}} = \frac{2.3RT}{F}[\log K_0 c_{\mathrm{b}} - f/2] \tag{14.19}$$

Setting $c_{\mathrm{b}} = 1$ and $f = 0$, we can define the standard potential for adsorption as

$$E_\theta^0 \equiv \frac{2.3RT}{F} \log K_0 \tag{14.20}$$

Substituting this into Eq. (14.19) we have

$$E_{\mathrm{p}} = E_\theta^0 + \frac{2.3RT}{F}[\log c_{\mathrm{b}} - f/2], \tag{14.21}$$

showing that the peak potential is shifted from its value for the Langmuir isotherm by an extent that is proportional to the Frumkin interaction parameter $f = r/RT$.

Cyclic voltammograms calculated for different values of the parameter f are shown in Figure 14.5. One should note that in the case of adsorbed intermediates discussed here, the peak current density is independent of concentration, while the peak potential depends on it through a Nernst-type equation. This is in contrast to similar equations for reaction of a bulk species, where j_{p} is proportional to the bulk concentration (Eqs. (14.9) and (14.11)), while E_{p} is independent of it, for a simple stoichiometry (Eqs. (14.4) and (14.10)). Also, the anodic and cathodic peaks for formation and removal of an adsorbed species occur at exactly the same potential, unlike the case of reaction of a bulk species.

Let us calculate the peak current density, for the purpose of comparing the current densities resulting from other processes, which may take place simultaneously. For the Langmuir isotherm, $C_{\phi,\max} = 2.3\ \mathrm{mF\ cm^{-2}}$, which gives rise to a peak current density of $j_{\phi,\mathrm{p}} = 0.23\ \mathrm{mA\ cm^{-2}}$ at $\nu = 0.10\ \mathrm{V\ s^{-1}}$, about 100 times the typical values of j_{dl} at the same sweep rate. Consequently, double-layer charging does not interfere seriously with measurement of the cyclic voltammogram. Moreover, since both j_{dl} and $j_{\phi,\mathrm{p}}$ depend linearly on sweep rate, their ratio is independent of it. Last, but not least, increasing the roughness factor is advantageous in this case,

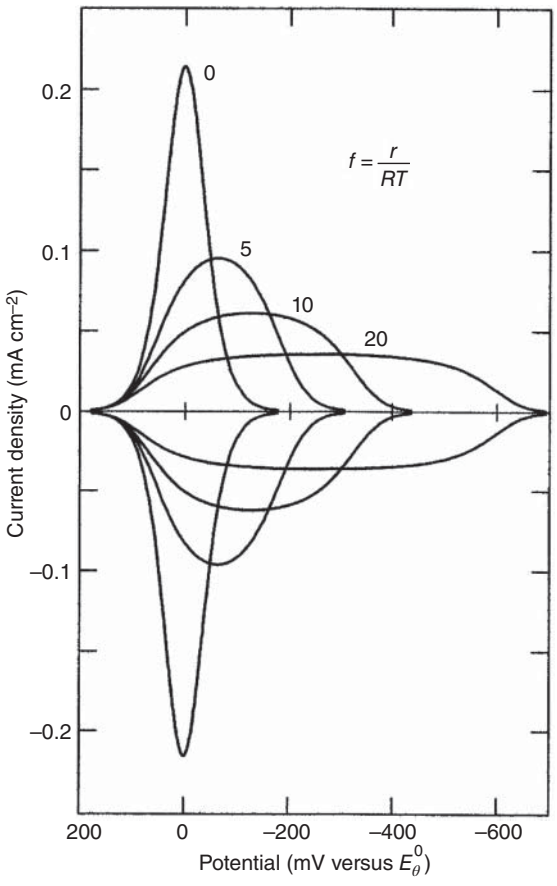


Figure 14.5 Simulated cyclic voltammogram for monolayer adsorption and desorption of species formed in a charge-transfer process. $v = 0.10\,\mathrm{V\,s^{-1}}$, $q_1 = 0.23\,\mathrm{mC\,cm^{-2}}$. The standard state is chosen as $\theta/(1-\theta) = 1$, $c_b = 1.0\,\mathrm{M}$ and $f = 0$.

for the same reason that it is a disadvantage in the case shown in Figure 14.3, since $j_{\phi,p}$ is proportional to the *real* surface area, while the diffusion-limited peak current density is proportional to the *geometrical* surface area (i.e. it is essentially independent of the roughness factor).

14.4.2 The High-Overpotential Region

In the reversible case, the rates of adsorption and desorption are so fast that the value of θ at any moment during the transient is equal to its equilibrium value. Having the fractional surface coverage controlled totally by the potential, through the appropriate adsorption isotherm, is equivalent to having the surface concentrations of reactants and products totally controlled by the Nernst equation, which is the assumption made for the reversible cyclic voltammetry. For the case of formation of atomic hydrogen on the surface of a platinum electrode, this assumption holds up to a sweep rate of about $1.0\,\mathrm{V\,s^{-1}}$. For the formation of $(OH)_{ads}$ on the same surface, it does not hold even for a slow sweep rate of $1.0\,\mathrm{mV\,s^{-1}}$. The transition from reversible to irreversible conditions is, of course, gradual.

Further Reading

1 Holze, R. (2015). *Cyclic Voltammetry – The Complete Guide*. Germany: Wiley-Blackwell.
2 Compton, R.D. and Banks, C.E. (2010). *Understanding Voltammetry*, 2nd edn. Imperial College Press.
3 Bard, A.J. and Faulkner, L.R. (2001). *Electrochemical Methods: Fundamentals and Applications*, 2nd edn. New York, USA: Wiley.
4 Gosser, D.K. (1993). *Cyclic Voltammetry Simulation and Analysis of Reaction Mechanisms*. Wiley-VCH.
5 Heinze, J. (1984). Cyclic voltammetry – "electrochemical spectroscopy". *Angew. Chem. Int. Ed.* 23: 831–847.
6 Laviron, E. (1979). General expression of the linear potential sweep voltammogram in the case of diffusionless electrochemical systems. *J. Electroanal. Chem.* 101: 19–28.
7 Macdonald, D.D. (1977). *Transient Techniques in Electrochemistry*. New York, USA: Plenum Press.
8 Gileadi, E., Kirowa-Eisner, E., and Penciner, J. (1975). *Interfacial Electrochemistry: An Experimental Approach*. Addison-Wesley.
9 Nicholson, R.S. (1965). Theory and application of cyclic voltammetry for measurement of electrode reaction kinetics. *Anal. Chem.* 37: 1351–1355.

15

Electrochemical Impedance Spectroscopy (EIS)

15.1 Introduction

The use of a phase-sensitive voltmeter for the study of the electrical response of the interphase can provide highly accurate measurements of the double-layer capacitance. However, this instrument has far more important uses in electrochemistry than just the measurement of capacitance. By combining a phase-sensitive voltmeter (also called a lock-in amplifier) with a variable frequency sine wave generator, one obtains an electrochemical impedance spectrometer. This makes it possible to probe the interphase over a wide range of frequencies and to record and analyze the data.

Electrochemical impedance spectroscopy (EIS) is an experimental AC technique that is based on the measurement of the complex impedance Z of an electrochemical system, which is either at steady state or at equilibrium. The impedance is measured as a function of the frequency f (Hertz, or cycles per second), or the angular frequency $\omega \equiv 2\pi f$ (rad s^{-1}), of an imposed sinusoidal perturbation of small amplitude. The amplitude of the applied potential is generally in the range of ± 5 to ± 25 mV at room temperature. The term "spectroscopy" stems from the use of frequency as a variable. Different processes at the surface of the electrode absorb electric energy at discrete frequencies, which causes time lag and measurable phase angle between the time-dependent excitation and the response signal. Such processes can be simulated with the aid of equivalent electrical circuits. The measurement is conducted within a large range of frequencies, say from millihertz to hundreds of kilohertz, or even microhertz to megahertz. The latter range of about 12 orders of magnitude is very high compared to other fields of spectroscopy. We recall, for instance, that visible light extends over a factor of just two in frequency. The range of frequencies that can be used in practice is limited by the electrochemical aspects of the system, not by instrumentation. Thus, measurements at very low frequencies take a long time, during which the interphase may change chemically. While it is technically possible to make measurements at, say, 10 μHz, this would take longer than a day to complete one cycle, and the changes in the interphase during the measurement at a single frequency could make the result meaningless. At the high-frequency end, stray capacitances and inductances combine with possible nonuniformity of current distribution at the electrode surface, to make the

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

results unreliable. For these reasons, it is recommended to limit measurements in the range of 10 mHz to 50 kHz, except when solid phases are tested.

The impedance of the system may be measured using various techniques, including frequency response analysis, phase-sensitive detection (PSD), and fast Fourier transform (FFT). Frequency response analyzers (FRAs) are instruments that determine the frequency response of a measured system. Their functioning is different from that of lock-in amplifiers. They are based on the correlation of the studied signal with the reference. The measured signal is multiplied by the sine and cosine of the reference signal of the same frequency and then integrated during one or more wave-periods. Such integration recovers the real and imaginary parts of the measured signal.

While the traditional way of measuring the impedance curve is point-by-point, i.e. by measuring the response to each individual sinusoidal perturbation with a frequency ω, some novel approaches have been developed based on the simultaneous imposition of a set of various sinusoidal harmonics, or noise, or a small-amplitude potential step, with subsequent Fourier and Laplace data analysis. The self-consistency of the measured spectra is tested with the use of the Kramers–Kronig transformations, whose violation testifies in favor of a non-steady-state character of the studied system (say, in corrosion).

The total impedance can be written in concise form as

$$Z(\omega) = \mathrm{Re}(Z) - i \times \mathrm{Im}(Z) \tag{15.1}$$

where $\mathrm{Re}(Z) \equiv Z'$ and $\mathrm{Im}(Z) \equiv Z''$ are the *real* and the *imaginary* parts of the impedance, respectively (sometimes called *resistance* and *reactance*, respectively), and $i \equiv \sqrt{-1}$ is the imaginary unit. It follows that the absolute value of the impedance vector (also known as the *impedance modulus*) is given by

$$|Z(\omega)| = [(\mathrm{Re}(Z))^2 + (\mathrm{Im}(Z))^2]^{1/2} \tag{15.2}$$

When a sinusoidal voltage signal

$$V(t) = V_{\mathrm{m}} \sin(\omega t) \tag{15.3}$$

is applied across the interphase (V_{m} is the voltage amplitude), the result is a sinusoidal current signal of the same frequency but displaced somewhat in time,

$$I(t) = I_{\mathrm{m}} \sin(\omega t + \phi) \tag{15.4}$$

namely, having a *phase shift* ϕ. Here we use the current I, instead of the current density j. The phase angle may be defined as

$$\tan(\phi) = \frac{\mathrm{Im}(Z)}{\mathrm{Re}(Z)} = \frac{1}{\omega RC} \tag{15.5}$$

and the real and imaginary parts as

$$\mathrm{Re}(Z) = |Z| \cos(\phi) \tag{15.6}$$

$$\mathrm{Im}(Z) = |Z| \sin(\phi) \tag{15.7}$$

The value $|Z|$ has units of resistance; it is the length of a vector obtained by addition of two perpendicular vectors, R and $1/\omega C$.

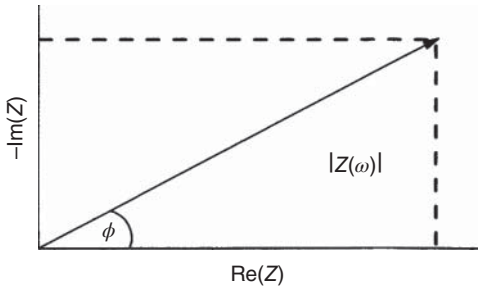


Figure 15.1 Vector representation of the impedance $Z(\omega)$ in the complex plane.

The AC analog to Ohm's law ($R = V/I$) is

$$Z(\omega) \equiv V(t)/I(\mathrm{t}) \tag{15.8}$$

Like resistance, impedance is a measure of the ability of a circuit to resist the flow of electric current.

In Figure 15.1 we can see the relations of the phase angle and the modulus to the real and imaginary components of the impedance.

For an ideally *polarizable* interphase with negligible solution resistance, the phase angle is −90°. For an ideally *nonpolarizable* interphase the phase angle is zero. Real systems do not behave ideally, of course. The actual phase angle will, therefore, be somewhere in between, and it will depend on frequency in most cases.

For a circuit containing both capacitors and resistors, the ratio between the applied voltage signal and the resulting current signal is the impedance $Z(\omega)$, which is a function of frequency. The impedance of a pure *resistor* is simply its resistance R, while the impedance of a pure *capacitor* is given by

$$Z_{\mathrm{C}} = -\frac{i}{\omega C} \tag{15.9}$$

For a capacitor and a resistor in *series* one has

$$Z(\omega) = R_{\mathrm{S}} - i/\omega C_{\mathrm{dl}} \tag{15.10}$$

while for a capacitor and resistor in *parallel* the impedance is given by

$$\frac{1}{Z(\omega)} = \frac{1}{R_{\mathrm{F}}} - \frac{\omega C_{\mathrm{dl}}}{i} \tag{15.11}$$

It is convenient to display the results of EIS in the *complex-plane impedance* representation. The x-axis on this plot is Re(Z), which is the Ohmic resistance, and the y-axis is −Im(Z), which, in the present case, is the capacitive impedance $-i/\omega C$. This is shown in Figure 15.2, which is a vector representation of $|Z|$ for both the series and parallel combinations.

As the frequency is *increased*, the capacitive impedance decreases, while the resistive impedance is unchanged. In the *series R–C circuit*, this makes the circuit behave more and more like a pure *resistor*, causing a decrease in the phase angle, as seen in Figure 15.2a. The impedance in this case is given by $Z = R - \mathrm{i}/\omega C$. The Nyquist graph in this case consists of a straight line perpendicular to the real axis, while the Bode modulus graph (not shown here) contains one breakpoint (following slope = −1). This point corresponds to the system characteristic frequency

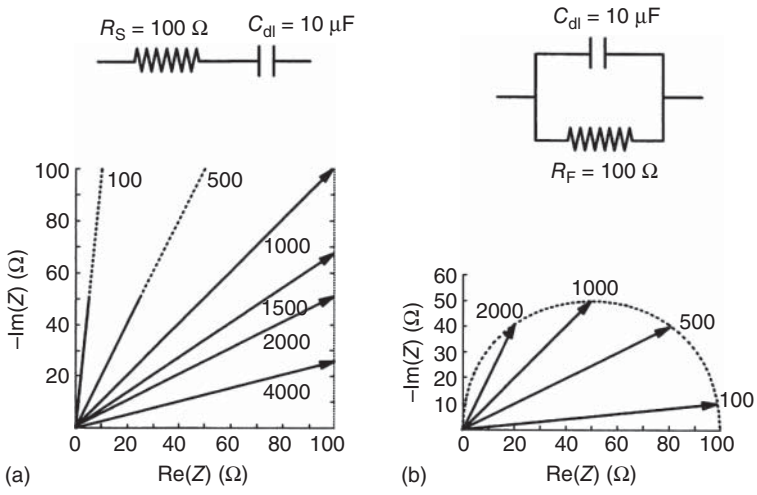


Figure 15.2 Vector representation of the impedance of (a) series and (b) parallel combination of a capacitor and a resistor, showing the variation of ϕ with ω (expressed in radians per second). The ends of the arrows show the absolute values of the vectors, except at the two lowest frequencies in the series combination.

$\omega = 1/RC$, or a time constant $\tau = RC$. The Bode phase-angle graph changes from 90° at low frequencies to 0° at high frequencies.

A *parallel R–C circuit* makes the circuit behave more and more like a *capacitor*, causing an increase in phase angle, as shown in Figure 15.2b. In this case, the total admittance is $Y(i\omega) = 1/R + i\omega C$, so that

$$Z(\omega) = \frac{1}{1/R + i\omega C} = \frac{R}{1 + i\omega RC} = \frac{R}{1 + \omega^2 R^2 C^2} - \frac{i\omega R^2 C}{1 + \omega^2 R^2 C^2} \qquad (15.12)$$

There are two impedance limits in this case: ($\omega = 0$, $Z = R$) and ($\omega \to \infty$, $Z = 0$). The Nyquist diagram will show a semicircle of radius $R/2$ with its center on the real axis, the frequency at the semicircle maximum being $\omega = 1/RC$. The circuit's characteristic breakpoint frequency (inverse of the characteristic time constant), as observed in the Bode modulus plot (not shown here), is the same as for the series R–C circuit. The complex plane admittance plot will show a straight line parallel to the imaginary axis, which is similar to the complex-plane impedance plot for the series R–C circuit.

Consider now a more realistic situation, in which both the series and the parallel resistance must be taken into account. The Randles equivalent circuit (see also Chapters 1 and 4) and the corresponding complex-plane impedance plot are shown in Figure 15.3. Let us derive the mathematical expression for the impedance applicable to this circuit, as a function of the angular frequency, ω:

$$Z(\omega) = R_S + \frac{1}{1/R_F - \omega C_{dl}/i} = R_S + \frac{R_F}{1 + i\omega C_{dl} R_F} \qquad (15.13)$$

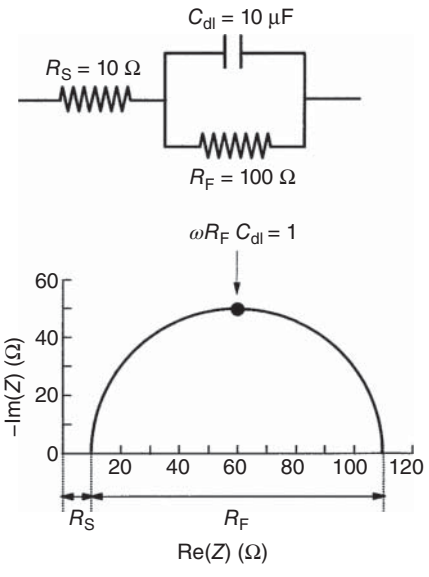


Figure 15.3 Complex-plane representation of the impedance of an interphase. The equivalent circuit in this case is known as the Randles equivalent circuit.

The following simple manipulation allows us to separate the real from the imaginary part of the impedance:

$$Z(\omega) = R_S + \left[\frac{R_F}{1 + i\omega C_{dl} R_F}\right] \times \left[\frac{1 - i\omega C_{dl} R_F}{1 - i\omega C_{dl} R_F}\right] \tag{15.14}$$

which leads to the expression

$$Z(\omega) = R_S + \frac{R_F}{1 + (\omega C_{dl} R_F)^2} - i \times \frac{\omega C_{dl} R_F}{1 + (\omega C_{dl} R_F)^2} \tag{15.15}$$

The result is a semicircle having a diameter equal to R_F, with its center on the x-axis and displaced from the origin of coordinates by $R_S + R_F/2$. Each point on the semicircle in Figure 15.3 represents a measurement at a given frequency. At the limit of high frequencies, the faradaic resistance is effectively shorted out by the double-layer capacitance, leaving the solution resistance R_S as the only measured quantity. At the limit of low frequency the opposite occurs: the capacitive impedance becomes very high, and one measures only the sum of the two resistors, $R_S + R_F$, in series. The same conclusion can be reached by substituting either high or low angular frequency in Eq. (15.13).

The double-layer capacitance can be obtained from this plot, since the maximum on the semicircle satisfies the equation

$$R_F \times C_{dl} \times \omega_{max} = 1 \tag{15.16}$$

When one studies an (almost) ideally polarizable interphase, such as the mercury electrode in pure deaerated acids, the equivalent circuit is a resistor R_S, and a capacitor C_{dl} in series. The high accuracy and resolution offered by modern instrumentation allows measurement in such cases in very dilute solutions or in poorly conducting, nonaqueous media, which could not have been performed in the early days of studying the mercury/electrolyte interphase.

15.2 Graphical Representations

The measured impedance data can be displayed in a number of different forms, each having its own coordinates:

(1) Plot of data for various frequencies as a set of points in the complex-impedance plane, $\mathrm{Re}(Z)$ and $-\mathrm{Im}(Z)$ being the x and y axes, respectively. This graph presentation is known as the complex-plane plot, *Nyquist* diagram, or Argand plot.
A *resistor* will be represented on the Nyquist plot by a single point on the real axis of this diagram; the impedance of a resistor is independent of frequency and has no imaginary component. With only a real impedance component, the current through a resistor stays in phase with the voltage across the resistor.
The impedance of an *inductor* increases as frequency increases. An inductor will be represented on the Nyquist plot by a vertical line coinciding with the imaginary axis. Since inductors have only an imaginary impedance component, the current through an inductor is phase-shifted −90° with respect to the voltage.
The impedance versus frequency behavior of a *capacitor* is opposite to that of an inductor; a capacitor's impedance decreases as the frequency is raised. Thus, a capacitor will be represented on the Nyquist plot by a vertical line coinciding with the imaginary axis. Capacitors also have only an imaginary impedance component. The current through a capacitor is phase-shifted −90° with respect to the voltage.
An *ideally polarizable* electrode behaves as an ideal capacitor because there is no charge transfer across the solution/electrode interphase. The impedance of an ideally polarizable electrode produces a straight line, perpendicular to the real axis of the Nyquist diagram.
The Nyquist plot has been commonly used in the electrochemical literature because it allows for an easy relation to the electrical model. However, the frequency itself is typically not shown on the Nyquist diagram, and this is one of the disadvantages of this form of presentation. This is sometimes corrected by marking the frequencies at which specific measurements were taken on the semicircle. *The diameter of the semicircle is equal to the faradaic resistance, and is independent of the capacitance.* As a result, plots measured for a fixed value of R_F but different values of C_{dl} cannot be distinguished in this type of presentation, even though corresponding points on the semicircle have been measured at different frequencies.

(2) Plots of $\log |Z|$ and ϕ versus $\log f$. These presentations are known as the *Bode modulus plot* and the *Bode phase-angle plot*, respectively. On the Bode modulus plot, a resistor will be represented by a horizontal line, whereas a capacitor will be represented by a line with a slope = −1. The Bode plots contain all the necessary information. That is why they are nowadays probably the most commonly used graphical presentations.

(3) Plots of $\mathrm{Re}(Z)$ and $\mathrm{Im}(Z)$ versus frequency (usually, $\log f$). This presentation is rarely used.

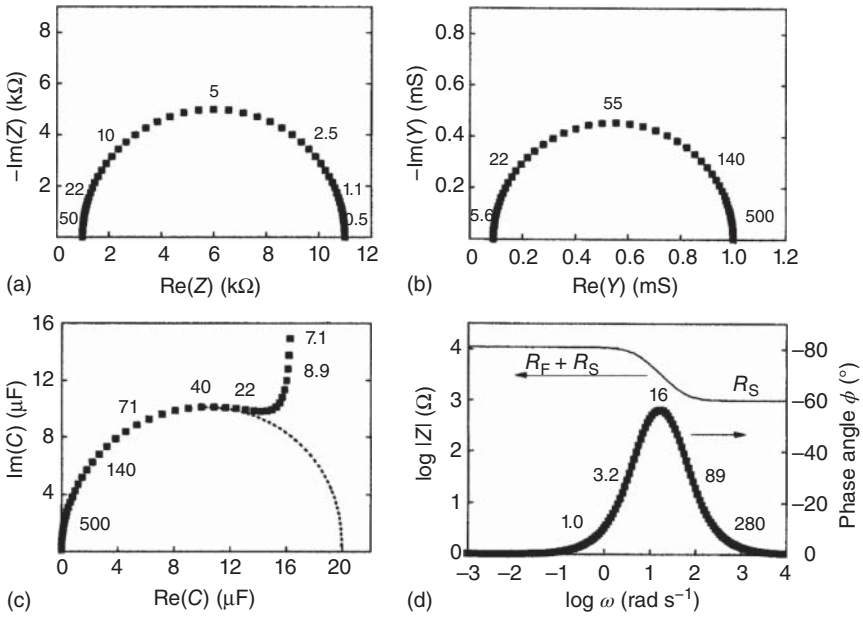


Figure 15.4 Comparison of (a) complex-plane impedance, (b) complex-plane admittance, (c) complex-plane capacitance, and (d) Bode modulus and phase angle plots for the same equivalent circuit. $R_F = 10\,k\Omega$, $R_S = 1\,k\Omega$, $C_{dl} = 20\,\mu F$. Values of ω (rad s^{-1}), at which some of the points were calculated, are shown.

(4) Recalculation of the impedance data to yield various other complex characteristics, most frequently the *admittance* $Y \equiv 1/Z$, and the complex capacitance $C \equiv Y/i\omega$. For the complex-plane admittance plot (see Figure 15.4b), the real and the imaginary parts of the admittance are defined as

$$\mathrm{Re}(Y) = \frac{\mathrm{Re}(Z)}{(|Z|)^2} \tag{15.17}$$

and

$$\mathrm{Im}(Y) = \frac{\mathrm{Im}(Z)}{(|Z|)^2} \tag{15.18}$$

For the complex-plane capacitance form (see Figure 15.4c)

$$\mathrm{Re}(C) = \frac{\mathrm{Im}(Z)}{(|Z|)^2}\frac{1}{\omega} \tag{15.19}$$

and

$$\mathrm{Im(C)} = \frac{\mathrm{Re(Z)}}{(|\mathrm{Z}|)^2}\frac{1}{\omega} \tag{15.20}$$

In Figure 15.4 above, different forms of presenting the response of the same circuit are compared. The values of the components constituting the equivalent circuit were chosen as $R_F = 10\,k\Omega$, $R_S = 1\,k\Omega$, and $C_{dl} = 20\,\mu F$. Each way of presentation has its own advantages and disadvantages. From the impedance plot, R_F and R_S can be read directly, and the double-layer capacitance can

be calculated, employing Eq. (15.16). The relevant time constant in this case is $\tau = R_F \times C_{dl} = 0.20\,s$; hence, $\omega = 5\,rad\,s^{-1} = 0.80\,Hz$. The complex-plane admittance plot yields the values of $1/R_F$ and $1/R_S$, but it is rarely used.

On the complex-plane capacitance plot, the semicircle results from the series combination of R_S and C_{dl} while the intercept with the real axis yields the value of C_{dl}. The vertical line is due to the parallel combination of R_F and C_{dl}. While the physical meanings of complex impedance and complex admittance are clear, that of the complex capacitance is questionable. Nevertheless, presenting the data as in Figure 15.4c has some merit, because the numerical value of C_{dl} is given directly as the diameter of the semicircle.

Plotting the same data in the Bode-type representation, one notes two points:

(a) Although the values of the two resistors are easily discerned, there is no region in which the circuit behaves as a pure capacitor. The slope never reaches a value of −1 in this figure, and ϕ never comes close to −90°, which one would expect for a pure capacitor.
(b) The phase angle is a more sensitive test of the capacitive or resistive behavior of the system than the plot of log $|Z(\omega)|$ versus log (ω). The detailed shape of these curves depends, of course, on the numerical values chosen for the various circuit elements. Had we used a value of $R_S = 10\,\Omega$ instead of 1 kΩ, the two horizontal lines in Figure 15.4d would have been much farther apart, and an (almost) pure capacitive behavior would have been observed in the intermediate region.

It would seem then that the complex-plane impedance plot is the best way of presenting the data, if one is mainly interested in the value R_F and its variation with time or potential. The complex-plane capacitance plot, on the other hand, brings out more directly the value of the capacitance and its variation with the different parameters of the experiment.

It should be remembered that the curves shown in Figure 15.4 are all simulated and therefore "ideal," in the sense that they follow exactly the equations used for the assumed equivalent circuit and the numerical values of its components. When conducting an experiment, the points are always scattered as a result of experimental error. Also, the frequency range over which reliable data can be collected does not necessarily correspond to the time constant of the system studied. Thus, for the case shown in Figure 15.4a, the semicircle can be constructed from measurements in the range of $1 \le \omega \le 20\,rad\,s^{-1}$. In Figure 15.4b one would have to use data in the range of about $10 \le \omega \le 200\,rad\,s^{-1}$ to evaluate the numerical values of the circuit elements. From the Bode magnitude plots, R_S can be evaluated from high-frequency measurements ($\omega \ge 100\,rad\,s^{-1}$), while R_F can be obtained from low-frequency data ($\omega \le 1\,rad\,s^{-1}$). The capacitance can be obtained *approximately* as $C_{dl} = 1/(\omega\,|Z|)$ at the inflection point (which coincides with the maximum on the Bode angle plot), but this is correct only if $\phi = -90°$, that is, if the circuit behaved as a capacitor at this frequency. In the present example, $\phi_{max} = -58°$ and the value of C_{dl} calculated in this manner is 18.4 μF, compared to the value of 20 μF used to calculate the curves.

When impedance spectroscopy was introduced to electrochemistry, there was some argument in the literature as to the best method of displaying EIS data.

Nowadays, the software provided with most commercial instruments allows us to display the data in all the ways shown in Figure 15.4 and choose the representation that best suits the particular purpose, making the above argument obsolete.

15.3 The Effect of Diffusion Limitation – The Warburg Impedance

So far in this chapter we have discussed only equivalent circuits that correspond to charge transfer, namely, the situation in which the faradaic resistance R_F is high and diffusion limitation is negligible. We might note in passing that EIS is inherently a *small-amplitude technique*, in which the AC component of the *j/E* relationship is in the linear region. This is most readily realized by maintaining the system at its OCP and applying a low-amplitude perturbation. However, the response to a small perturbation can be linear, even if the system is in the nonlinear region. The faradaic resistance measured by EIS is the *differential* resistance, defined as

$$R_F \equiv (\partial\eta/\partial j)_\mu \tag{15.21}$$

In the high-overpotential region, j is an exponential function of the overpotential; hence, the differential faradaic resistance also depends exponentially on potential:

$$\eta = b \log j/j_0 \tag{15.22}$$

Combining the last two equations, one has, for the high-overpotential region,

$$R_F = \frac{b}{j} \tag{15.23}$$

In the low-overpotential region, where the current is linearly related to the overpotential,

$$\frac{j}{j_0} = \frac{nF}{RT}\eta \tag{15.24}$$

Hence, in this region the faradaic resistance is given by

$$R_F = \frac{\eta}{j} = \frac{1}{j_0}\frac{RT}{nF} \tag{15.25}$$

It is very important to note the difference between the two expressions for the differential faradaic resistance, given in Eqs. (15.23) and (15.25). At high overpotential, R_F is inversely proportional to the current density, but it yields no information concerning the exchange current density (and, hence, of the heterogeneous rate constant). For example, it is well known that j_0 for electrodeposition of Cu is about 3 orders of magnitude higher than that for Ni. Nevertheless, if EIS is used to measure R_F at the same current density, the same value will be observed in both cases, unless the Tafel slope is different. However, this change could be a factor of 2 or 3 at most, while the exchange current density could vary over several orders of magnitude. In contrast, when EIS is applied to a system at its

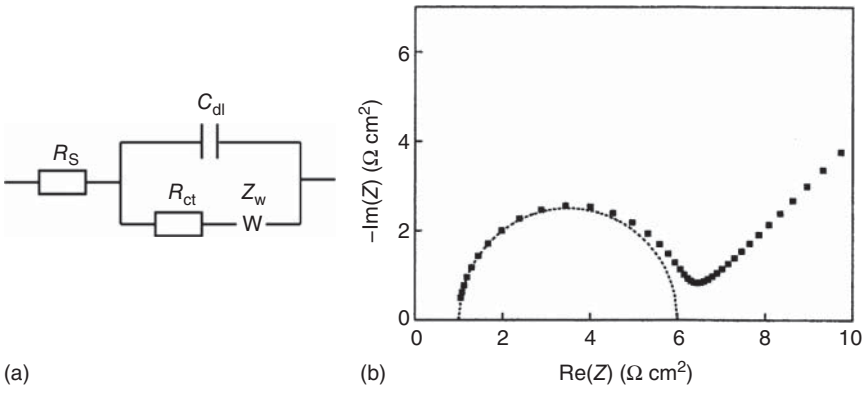


Figure 15.5 (a) Modified Randles equivalent circuit due to diffusion limitation. (b) Complex-plane impedance plot for an equivalent circuit with diffusion limitation at low frequencies. The values chosen for this simulation were $R_F = 5\,\Omega\,\mathrm{cm}^2$, $R_S = 1\,\Omega\,\mathrm{cm}^2$, $C_{dl} = 20\,\mu\mathrm{F\,cm}^{-2}$, $c_{b,Ox} = c_{b,Red} = 10\,\mathrm{mM}$, $D_{Ox} = D_{Red} = 1 \times 10^{-5}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$, $\sigma = 12\,\Omega\,\mathrm{cm}^2\,\mathrm{s}^{-1/2}$, $j_0 = 5\,\mathrm{mA\,cm}^{-2}$, $k_{s,h} = 5 \times 10^{-3}\,\mathrm{cm\,s}^{-1}$. Frequency range: $1 \times 10^{-3} \le \omega \le 1 \times 10^3\,\mathrm{rad\,s}^{-1}$.

reversible potential, or at the OCP, Eq. (15.25) is applicable, and the value of j_0 can be obtained directly.

Turning now to the case in which diffusion control must be considered (e.g. under conditions of concentration polarization, when surface layer is formed, or when top coating is present), the equivalent circuit takes the form shown in Figure 15.5a, in which the symbol "W" represents the so-called *Warburg impedance*, which accounts for diffusion limitation. The diffusion equations have been solved for the low-amplitude sine wave perturbation. It is found that the Warburg impedance is given by

$$Z_W = \frac{\sigma}{\omega^{1/2}} - i\frac{\sigma}{\omega^{1/2}} \tag{15.26}$$

where the parameter σ, also known as the Warburg impedance coefficient, is defined as

$$\sigma = \frac{RT}{(nF)^2}\frac{1}{\sqrt{2}}\left[\frac{1}{c_{b,Ox} \times D_{Ox}^{1/2}} + \frac{1}{c_{b,Red} \times D_{Red}^{1/2}}\right] \tag{15.27}$$

This equation can be simplified if it is assumed that the two diffusion coefficients are equal, and the bulk concentrations of the oxidized and the reduced species are also equal, yielding

$$\sigma = \frac{RT}{(nF)^2}\frac{2^{1/2}}{c_b \times D^{1/2}} \tag{15.28}$$

If the two concentrations differ widely, Eq. (15.27) shows that the lower concentration determines the value of σ.

One should note that the real and the imaginary parts of the Warburg impedance in Eq. (15.26) depend on frequency in the same way. Therefore, the phase shift generated by the Warburg impedance is independent of frequency.

Plotted in the complex-plane impedance format, this leads to a straight line with a slope of unity (a phase shift of 45°), as shown in Figure 15.5.

Equations (15.26)–(15.28) were derived for conditions of semi-infinite linear diffusion. Thus, measurements should be conducted in quiescent solutions. In addition, if there is a thin film on the surface through which diffusion occurs, such as the solid electrolyte interphase (SEI) formed in some nonaqueous batteries, the condition of semi-infinite linear diffusion no longer applies, and the part corresponding to diffusion in Figure 15.5 will be different.

We note that the Warburg impedance, which is proportional to $\omega^{-1/2}$, is in series with the faradaic resistance, R_F, as seen in Figure 15.5. At high frequencies, one obtains the usual semicircle, whereas the Warburg impedance becomes predominant at low frequencies. The frequency at which the transition occurs depends on the concentrations of reactants and products (which determine the value of σ), on the exchange current density, and on the overpotential, which determine the value of the faradaic resistance, R_F.

Four complex-plane impedance plots, calculated for the concentrations of 100, 10, 1, and 0.1 mM, respectively, are shown in Figure 15.6a–d. Note that R_F, which is proportional to $1/j_0$, also depends on concentration; therefore, the scale of both Im(Z) and Re(Z) must be changed from Figure 15.6a–d. In calculating the curves of this figure, it was assumed that j_0 is a linear function of concentration, which is not always the case. At the highest concentration shown here, a semicircle characteristic of a charge-transfer-limited process is clearly seen, with diffusion limitation becoming important only at low frequencies. As the concentration is decreased, diffusion limitation becomes gradually more important. In a 1 mM solution, the initial part of the semicircle is barely seen, while in a 0.1 mM solution the process is mostly diffusion controlled. The values of R_F chosen here correspond to a moderately slow reaction, having a standard heterogeneous rate constant of $k_{s,h} = 2.7 \times 10^{-3}$ cm s^{-1}.

It is important to realize that the plots we have shown above are all based on simulated data and serve the purpose of illustrating the principles involved. The results of real experiments are rarely as simple and easy to interpret. This is caused by two types of factors: (i) the use of EIS is straightforward only when the behavior is simple and can be modeled by an equivalent circuit that involves just a few adjustable parameters, and (ii) the real picture found experimentally is rarely as simple as that shown by simulation. When complex shapes of the impedance spectrum are observed, one must develop a physical model and choose a corresponding equivalent circuit that can simulate the observed results. However, when fitting the model requires many adjustable parameters, the validity of the model becomes less reliable. In such cases, an independent experiment, such as considering the dependence of different parts of the spectrum on factors such as potential or current density, the concentration of the electroactive components, as well as the supporting electrolyte, can be useful. However, this approach weakens the usefulness of EIS in the study of electrode kinetics – rather than employing this technique to evaluate the physical model, one may end up using an assumed model to explain the complex impedance spectrum.

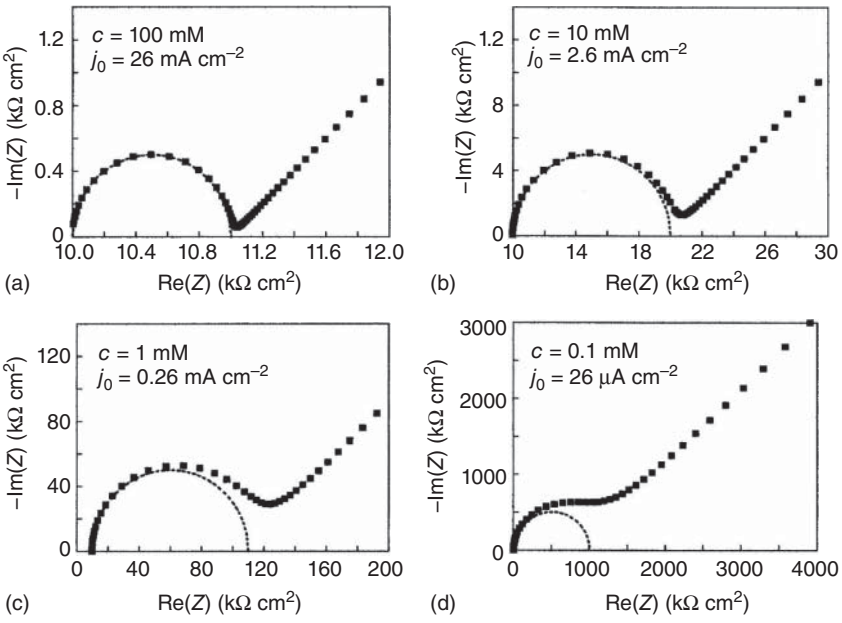


Figure 15.6 Complex-plane impedance plots showing the gradual change from charge-transfer to mass-transport control with decreasing concentration. $R_S = 10\,\Omega\,cm^2$, $C_{dl} = 20\,\mu F\,cm^{-2}$, $j_0 = 26, 2.6, 0.26, 0.026\,mA\,cm^{-2}$ for (a), (b), (c), and (d), respectively. Frequency range: $1 \times 10^{-3} \leq \omega \leq 1 \times 10^3$ rad s^{-1}.

15.4 Advantages, Disadvantages, and Applications of EIS

The advantages of EIS include the following: (i) The absence of any significant perturbation to the system. (ii) Its applicability to the assessment of low-conductivity media such as polymers. (iii) The existence of a frequency component that may provide mechanistic information. Furthermore, EIS may be able to distinguish between two or more electrochemical reactions taking place. (iv) EIS can test components within an assembled device using the device's own electrodes. (v) EIS can identify diffusion-limited reactions, e.g. diffusion through a passive film.

However, EIS also has some significant drawbacks: (i) At low frequencies, where underfilm corrosion reactions, for example, are probed, experimental difficulties and time constraints can complicate precise determination of the charge-transfer resistance R_{ct} and the double-layer capacitance C_{dl}. (ii) The Warburg impedance may exhibit a masking effect. (iii) Inductive loops or new time constants are sometimes observed at low frequency. Such inductive loops are generally ascribed to the existence of relaxation processes of adsorbed species (e.g. relaxation of coverage due an adsorbed intermediate), to surface area or salt film property modulation, etc. (iv) Owing to the long measurement time required, both the corrosion rate and corrosion potential can change during the experiment. (v) There has been no evidence that R_{ct} can be equated to R_p and be used in the Stern–Geary equation to calculate the underfilm corrosion rate.

(vi) In-depth mechanistic and kinetic information on the corroding interface can easily be overlooked in EIS by modeling the system with an oversimplified equivalent circuit.

EIS is a mature technique. It is one of the most widely employed techniques for studying the corrosion protection by paints and organic coatings, for predicting the lifetime of coating/metal systems, and for determining the effectiveness of corrosion inhibitors. It is also used to assess interfacial reactions; adsorption to and desorption from an electrode surface; chemical reactions coupled with faradaic processes; conducting polymers; electrochemical synthesis of materials; ions mobility in energy storage and conversion devices such as batteries, supercapacitors, and fuel cells; semiconductors; kinetics of electrocatalytic reactions; label-free detection sensors; etc.

Here, we give as an example the use of EIS for assessing the development of corrosion of ferromagnetic cobalt (in the forms of foil and e-beam deposit on silicon wafer), either bare or coated with an acrylic polymer, in 0.5 M NaCl electrolyte solution. The potential was maintained at the OCP during the experiment, with a superimposed perturbation of 20 mV for coated samples and 5 mV for bare samples. The frequency range was 20 mHz to 50 kHz. The most common equivalent circuit employed for polymer-coated systems is presented in Figure 15.7a. In this figure, R_f and C_f are the film (coating) resistance and capacitance. Diffusion processes are represented by the Warburg impedance Z_W, which can be related to blockage of oxygen diffusion paths by corrosion products. For a non-degraded polymer-coated metal system, before the interface becomes active, only the external circuit (R_S, R_f, and C_f) will be observed experimentally. With time, however, the inner circuit (R_{ct}, C_{dl}, and Z_W) may also become active, and this is the source of a second time constant. The appearance of the latter has been linked with system breakdown. The two time constants represent the film (τ_f) at higher frequencies and the metal (τ_m) at lower frequencies. In order to determine the rate of corrosion and, hence, the mass loss, the R_{ct} (or R_p) value needs to be detected. While in the simplest situation, R_{ct} and R_p may be equivalent, these values are not necessarily the same.

Figure 15.7b–d presents the EIS Bode and Nyquist plots for bare cobalt with increasing times of exposure to the electrolyte. The polarization resistance was measured as the difference between the two plateaus in the Bode modulus plot (Figure 15.7d). Similar values were obtained by measuring the diameter of the depressed semicircle in the Nyquist plot. Sweeps performed at times longer than 69 hours of immersion yielded spectra that could not be deconvolved for analysis. Such behavior may result from the significantly reduced volume of cobalt and the exposure of the silicon wafer beneath the cobalt film.

In the case of polymer-coated cobalt, EIS analysis could only be made after some immersion time, when the cobalt beneath the polymer was clearly exposed. This was reflected by the appearance of a second semicircle at low frequencies in the Nyquist plot. EIS analysis was also not possible after long immersion times, likely due to masking effects of diffusion through corrosion products, which had formed at the bottom of defects in the coating and/or exposure of the silicon substrate. Selected Bode and Nyquist plots for the coated samples are shown in Figure 15.7e,f.

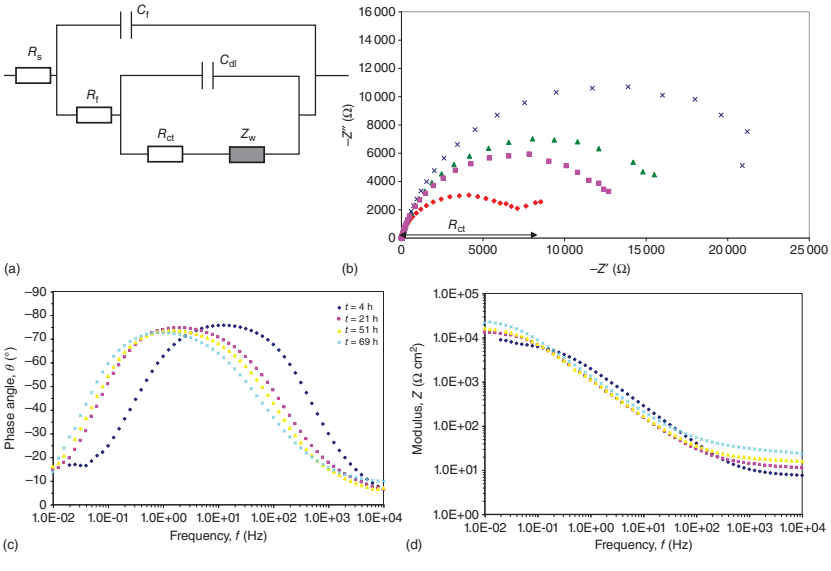


Figure 15.7 (a) The equivalent circuit frequently used to represent polymer-coated metals. In this figure, R_f and C_f are the coating resistance and capacitance, respectively. Source: Adapted from Mitton et al., *J. Electrochem. Soc.* **149** (2002) B265. Nyquist (b) and Bode (c,d) EIS plots of bare cobalt after 4, 21, 51, and 69 hours of immersion. Sweeps performed at times longer than 69 hours of immersion yielded spectra, which could not be deconvolved for analysis. (e,f) EIS Nyquist plots for polymer-coated cobalt after 6, 93, 169, and 302 hours of immersion. Only the data after 93 and 169 hours could be used for the determination of R_{ct}. Note the different scales of the Z' and Z'' axes in (e) and (f). Source: Adapted from Eliaz et al., *Mater. Technol.* **16** (2001) 90.

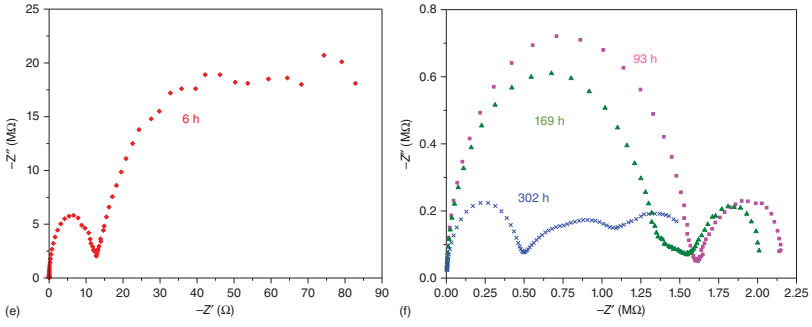


Figure 15.7 (*Continued*)

How can we determine the cumulative mass loss due to corrosion based on EIS measurements? First, the corrosion rate, CR (g cm^{-2} s^{-1}), should be determined:

$$\mathrm{CR} = \frac{a_\mathrm{m} \cdot j_\mathrm{corr}}{nF} \tag{15.29}$$

where a_m is the atomic mass of the corroding metal (g mol^{-1}), j_corr is the corrosion current density (A cm^{-2}), n the number of electrons transferred in the reaction (dimensionless), and F the Faraday constant (C mol^{-1}).
Substituting the Stern–Geary relation for j_corr,

$$\mathrm{CR} = \frac{a_\mathrm{m}}{nFAR_\mathrm{p}} \cdot \frac{b_\mathrm{a}|b_\mathrm{c}|}{2.3(b_\mathrm{a} + |b_\mathrm{c}|)} \tag{15.30}$$

where A is the exposed surface area (cm^2), R_p the polarization resistance (Ω cm^2), and b_a and b_c are the anodic (oxidation) and cathodic (reduction) Tafel slopes (V decade^{-1}).

The cumulative mass loss is thus

$$\Delta m = \int_{t_0}^{t_n} (\mathrm{CR}) A \, \mathrm{d}t \tag{15.31}$$

The capacitance of an organic coating, C_f, is an important measure of coating failure. It is defined by

$$C_\mathrm{f} = \frac{\varepsilon \varepsilon_0 A}{d} \tag{15.32}$$

where ε is the dielectric constant of the coating, ε_0 is the permittivity of free space (electric constant), A is the surface area of the working electrode exposed to the solution, and d is the coating thickness. Because most organic coatings are relatively thick, the coating capacitance is typically rather low, at around 1 nF cm^{-2}.

EIS can be used to monitor the absorption of water in organic coatings. The impedance from the Bode plot decreases when the coating absorbs water, but maintains a linear relationship with frequency. The slope does not change, and the phase angle remains −90°. Water has a higher dielectric constant (78.5 at 25 °C) than that of the typical organic coating (2–5). Therefore, the capacitance of the coating with absorbed water is higher than that of the dry coating (see Eq. (15.32)). The coating may also swell during this stage, thus resulting in decrease of the capacitance. Fortunately, because of the high dielectric constant of water, the effect of water absorption usually overwhelms the effect of swelling, and the capacitance increases. The initial effect of water absorption is to change the "value" of the capacitor, so the shape of the Nyquist plot does not change. The Brasher–Kingsbury equation provides a convenient estimate of water uptake in the coating:

$$\text{Volume fraction of } \mathrm{H_2O} = \frac{\log(C_{\mathrm{f,t}}/C_{\mathrm{f},0})}{\log \varepsilon_\mathrm{w}} \tag{15.33}$$

where $C_{\mathrm{f,t}}$ is the coating's capacitance at time t, $C_{\mathrm{f},0}$ is the initial coating's capacitance, and ε_w is the dielectric constant of water. The coating's capacitance can be easily measured at any time by fitting the equivalent circuit to the EIS data.

Further Reading

1 Lasia, A. (2014). *Electrochemical Impedance Spectroscopy and its Applications*. New York, USA: Springer-Verlag.
2 Bard, A.J., Inzelt, G., and Scholz, F. (ed.) (2012). *Electrochemical Dictionary*. Berlin Heidelberg: Springer-Verlag.
3 Orazem, M.E. and Tribollet, B. (2008). *Electrochemical Impedance Spectroscopy*. New York, USA: Wiley.
4 Mitton, D.B., Wallace, S.L., Cantini, N.J. et al. (2002). The correlation between substrate mass loss and electrochemical impedance spectroscopy data for a polymer-coated metal. *J. Electrochem. Soc.* 149: B265–B271.
5 Turgoose, S. and Cottis, R. (1999). *Electrochemical Impedance and Noise*. NACE International.
6 Bonora, P.L., Deflorian, F., and Fedrizzi, L. (1996). Electrochemical impedance spectroscopy as a tool for investigating underpaint corrosion. *Electrochim. Acta* 41: 1073–1082.
7 McIntyre, J.M. and Pham, H.Q. (1996). Electrochemical impedance spectroscopy: a tool for organic coatings optimizations. *Prog. Org. Coat.* 27: 201–207.
8 Mansfeld, F. (1995). Use of electrochemical impedance spectroscopy for the study of corrosion protection by polymer coatings. *J. Appl. Electrochem.* 25: 187–202.
9 Macdonald, J.R. (1987). *Impedance Spectroscopy: Emphasizing Solid Materials and Systems*. New York, USA: Wiley.
10 Brasher, D.M. and Kingsbury, A.H. (1954). Electrical measurements in the study of immersed paint coatings on metal. I. Comparison between capacitance and gravimetric methods of estimating water-uptake. *J. Appl. Chem.* 4: 62–72.
11 Randles, J.E.B. (1947). Kinetics of rapid electrode reactions. *Discuss. Faraday Soc.* 1: 11–19.

16

The Electrochemical Quartz Crystal Microbalance (EQCM)

16.1 Fundamental Properties of the EQCM

16.1.1 Introduction

In 1880, Jacques and Pierre Curie discovered the *piezoelectric effect*, according to which, when mechanical stress is applied to the surfaces of various crystals (changing their dimensions), a potential difference is generated across these crystals, whose magnitude is proportional to the applied stress. Later, they discovered also the *inverse piezoelectric effect*, according to which, the application of potential across the two sides of the crystal will result in a corresponding mechanical strain (i.e. dimensional changes). When the potential applied is a periodic signal (typically, a sine wave), the crystal vibrates at the frequency of the applied signal. These piezoelectric properties are the operational basis of the quartz crystal microbalance (QCM), which was first used in a sensing mode when Sauerbrey reported a linear relationship between the decrease in the resonance frequency, f_0, of an oscillating quartz crystal and the bound elastic mass of deposited metal.

When a quartz crystal (or any other solid material) vibrates, there is always a resonance frequency where it oscillates with minimum impedance (that is, maximum admittance). The resonance frequency depends on the dimensions and on the properties of the vibrating crystal, mostly the density and the shear modulus. A quartz crystal can be made to oscillate at other frequencies. However, as the distance (on the scale of frequency) from the resonance frequency increases, the admittance decreases, until the vibration can no longer be detected. This is the basis for the analysis of the so-called (mechanical) admittance spectrum of the QCM, which will be discussed below.

An important property of a vibrating quartz crystal is that its resonance frequency changes when its mass is changed, either by adding or removing mass or by interaction with the medium in which the crystal is immersed. Since frequency, and changes in it, can be measured with very high accuracy, a device based on a quartz crystal, coated on both sides with a thin layer of metal, can serve as a mass-sensitive detector, namely, a QCM.

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

16.1.2 The EQCM

Figure 16.1a shows the schematics of the EQCM. A thin disk sliced from single crystals of α-quartz is sandwiched between two metal electrodes that are vapor deposited on the two sides of the crystal (Figure 16.1b,c). Gold electrodes have been the most commonly used in QCM studies because of the ease with which Au can be evaporated. However, Cu, Ni, Pt, Ti, and other metals have also been employed. To provide a good adhesion of the metal film to the quartz plate, a very thin chromium layer is usually evaporated. When an alternating electric field is applied to the electrodes, the quartz crystal starts to oscillate. The result of the vibration motion of the quartz crystal is the establishment of a transverse acoustic wave that propagates across the thickness of the crystal, reflecting back into the crystal at the opposite surface. A standing-wave condition can be established in the quartz resonator when the acoustic wavelength is equal to twice the combined thickness of the crystal and electrodes. Thus, the resonance frequency of the quartz crystal prior to the mass change can be related to the thickness of the crystal, t_q, by the following relation:

$$f_0 = \frac{\nu_{tr}}{2t_q} \tag{16.1}$$

where ν_{tr} is the transverse velocity of sound (in AT-cut quartz, $\nu_{tr} = 3.34 \times 10^4\ \mathrm{m\,s^{-1}}$). This relation assumes that the thickness of the electrodes is small in comparison with that of quartz, so that the velocities of sound in the electrodes and in the quartz are identical. The acoustic velocity depends on the shear modulus and density of the crystal.

Quartz is the most stable form of silica (SiO_2). In QCM, α-quartz crystals are employed, because of their superior mechanical and piezoelectric properties. Quartz is a highly anisotropic material. The cut angle with respect to crystal orientation determines the mode of oscillation and other elastic properties. The AT-cut crystal, which is the most commonly used for QCM applications, is fabricated by slicing through a quartz rod with a cut angle of 35°10′ with respect to the optical axis, as shown in Figure 16.1d; this crystal exhibits shear displacement perpendicular to the resonator surface. AT marks that it was the first (A) temperature compensated (T) cut to be used. Other cuts are LC (linear cut) which can be used as a temperature sensor; BT (another temperature-compensated) cut, but with a stress response factor opposite to that of the AT-cut; and SC (stress-compensated) cut, which is insensitive to stress in the deposited material. The advantage of the AT-cut quartz crystal is that it has the smallest frequency shift with temperature around room temperature. It has the following properties: shear modulus $\mu_q = 2.947 \times 10^{10}\ \mathrm{N\,m^{-2}}$, density $\rho_q = 2.648\ \mathrm{g\,cm^{-3}}$.

The quartz crystal behaves like a circuit composed of an inductor, capacitor, and resistor in series, with a precise resonance frequency (Figure 16.1e). Adding capacitance across a crystal will cause the parallel resonance to shift downward. This can be used to adjust the frequency at which a crystal oscillator oscillates. Crystal manufacturers normally cut and trim their crystals to thickness yielding a specified resonance frequency and a known "load" capacitance added to

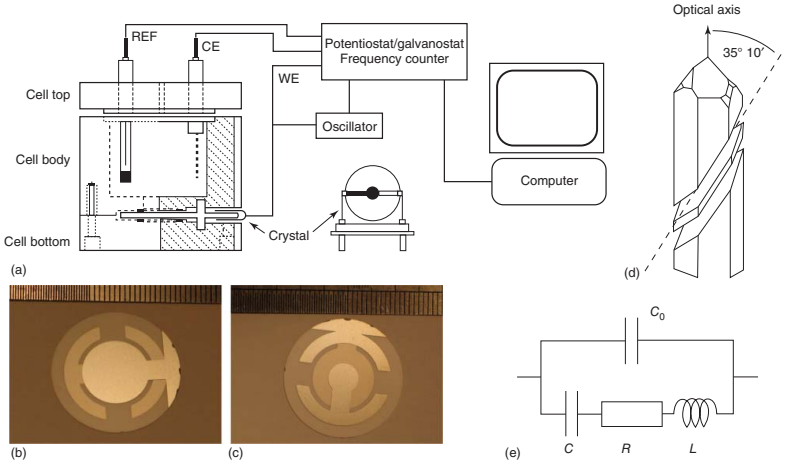


Figure 16.1 (a) Schematics of an EQCM, which allows applying a potential on the upper metal electrode, thereby creating an electrochemical cell. This enables the measurement of the change of mass in electrochemical reactions. The upper side (b) and bottom side (c) of a metal-coated quartz disk. The upper side of the electrode is sealed with an O-ring, so that the bottom side remains dry and provides a two-point short-circuit to the mass sensor. (d) The AT-cut of a quartz crystal. (e) The equivalent circuit for a quartz crystal in an oscillator.

the crystal. Typical operating frequencies of the QCM lie within the range of 5–10 MHz (determined in vacuum or in H_2 or He at ambient pressure), which provides for mass detection limits approaching 1 ng cm^{-2}. From a sensitivity standpoint, a high base frequency is desirable, since the mass sensitivity is proportional to f_0^2. For an AT-cut 10 MHz quartz, the thickness is about 170 μm. The quartz diameter is typically around 1.25 cm.

The QCM was introduced in 1959 by Sauerbrey, who derived the following equation to describe the correlation between the measured frequency change, Δf, and the elastic mass change per unit of the piezoelectrically active area, Δm:

$$\Delta f = -C_m \Delta m \tag{16.2}$$

where C_m is a characteristic constant given by

$$C_m = 2f_0^2(\mu_q \rho_q)^{-1/2} \tag{16.3}$$

For $f_0 = 6$ MHz (which we shall use for sample calculations in this chapter), Eq. (16.3) yields a numerical value of

$$C_m = 8.15 \times 10^7 \ \text{Hz cm}^2 \ \text{g}^{-1} \tag{16.4}$$

Thus, a change of frequency of −1.0 Hz is generated by an added mass of 12.3 ng cm^{-2}. The introduction of the QCM represented a significant increase in the sensitivity of weighing materials in vacuum or in the gas phase at ambient pressure, and was soon applied as an online method for measuring and controlling the rate of deposition of very thin films, for example, during sputtering or physical and chemical vapor deposition.

The following simple calculation demonstrates the high sensitivity of the EQCM. A monolayer of metal deposited on the surface amounts to about 2 nmol cm^{-2}. Hence, the mass of a single atomic layer of silver, for example, is about 220 ng cm^{-2}, corresponding to a frequency shift of about 18 Hz. Already when invented in 1959, it was possible to measure a frequency change of 1 Hz, and presently available instrumentation can take us to a resolution of 0.05 Hz, corresponding to 0.6 ng cm^{-2}, allowing the detection of changes in the amount of silver deposited of less than 0.3% of a monolayer.

The Sauerbrey equation is accurate as long as the thickness of the film added to (or removed from) the QCM is less than 2% of the thickness of the quartz crystal. Equations (16.2)–(16.4) look simple and straightforward, allowing the determination of mass with a resolution of the order of a few nanogram per square centimeter.[1] There seems to be no need to calibrate such a device, because the value of C_m given in Eq. (16.4) is based only on the properties of the quartz crystal. Nevertheless, it should be remembered that there are some tacit assumptions that must be fulfilled for Eq. (16.2) to be applicable. First, the deposition of any material must be uniform all over the effective area of the crystal. Secondly, measurements must be maintained within the dynamic range of the device, which is about ±100 kHz. This corresponds to an added mass of 1.2 mg cm^{-2}, or to a thickness 1.2 μm of silver. Next, it is assumed that the adsorbate is rigidly attached to

1 Indeed, one is tempted to call it a quartz crystal nanobalance, but we shall maintain the commonly used name of QCM, in order to avoid confusion.

the surface and that the surface roughness does not change during the experiment. This becomes increasingly important as the pressure of the gas increases, and could become critical when the EQCM is immersed in a liquid.

Control of the temperature is also important. As aforementioned, the crystals most commonly used are of the AT-cut type. This choice is made because it has the lowest dependence of the resonance frequency on temperature. Nevertheless, there is still some temperature dependence amounting to

$$\left(\frac{\mathrm{d}f_0}{\mathrm{d}T}\right) = 2.4 \ \mathrm{Hz\,^{\circ}C^{-1}} \tag{16.5}$$

Thus, in order to maintain a resolution of $\pm 0.6\,\mathrm{ng\,cm^{-2}}$ in the determination of the added mass, one would have to maintain the temperature constant at about $\pm 0.02\,^{\circ}\mathrm{C}$.

The total pressure must also be maintained constant, because the resonance frequency depends on pressure, following the equation

$$\Delta f_{\mathrm{P}} = [1.06 \times 10^{-6} f_0]\Delta p \tag{16.6}$$

where the pressure is expressed in atmospheres.

16.1.3 The Effect of Viscosity

When the EQCM was first introduced, it was not clear if this device could operate in condensed media, and specifically in aqueous solutions. Given the tacit assumptions that must be fulfilled for Eq. (16.2) to be applicable, the Sauerbrey equation is only strictly applicable to uniform, rigid, thin-film deposits. Consequently, the QCM was for many years just regarded as a gas-phase mass detector. At the beginning of 1980s scientists realized that a quartz crystal could be excited to a stable oscillation when it was completely immersed in a liquid. Much of the pioneering work in liquid-phase QCM measurements was done by Kanazawa and coworkers, who showed that the interaction of vibrating EQCM in a liquid leads to an added term in the change in resonance frequency, which depends primarily on the square root of the product of the viscosity and density of the fluid, namely,

$$\Delta f_{\eta} = -C_{\eta}(\rho_{\mathrm{fl}}\,\eta_{\mathrm{fl}})^{1/2} \tag{16.7}$$

where ρ_{fl} and η_{fl} refer to the density and viscosity of the fluid, respectively, and the constant determining the effect of viscosity and density is given by

$$C_{\eta} = f_0^{3/2}(\pi\,\mu_{\mathrm{q}}\rho_{\mathrm{q}})^{-1/2} \tag{16.8}$$

It is noted that both mass and viscosity represent added loads; therefore, they both decrease the resonance frequency. Considering that the density of liquids is usually 2 to 3 orders of magnitude higher than that of gases at ambient pressures, it is obvious why Eq. (16.7) assumes major importance when the EQCM is immersed in a liquid.

Equation (16.7) is also important when considering the effect of temperature on the shift of the resonance frequency, because the viscosity of liquids is temperature dependent, and this effect happens to be particularly high for aqueous

solutions. Thus, the effect of temperature on the EQCM immersed in a liquid could be higher than that indicated by Eq. (16.5).

Once it has been confirmed that the EQCM could operate in contact with liquids, its use in electrochemistry for the study of the metal/solution interphase became obvious and widespread.

16.1.4 Immersion in a Liquid

When the EQCM is immersed in a liquid, the resonance frequency is decreased substantially. Based on Eq. (16.7) above, it should decrease by about 1 kHz, but in practice it is usually found to decrease by 2–3 kHz. This does not imply that Eq. (16.7) is incorrect. It only shows that there must have been some tacit assumption which, alas, does not apply to real systems. In the present case, we failed to note that Eq. (16.7) applies to an ideally flat surface, while almost all real surfaces are rough to some extent. Even when a surface is highly polished, there is some roughness remaining. It is hard to describe the extent or degree of roughness, because roughness is inherently irregular. On the other hand, it is tentatively clear that a rough vibrating surface would interact with the liquid in which it is immersed much more strongly than an ideally flat surface. While it is hard to prepare an ideally smooth surface, it is easy to prepare surfaces of varying roughness. One can create increasingly rough surfaces by plating a metal at current densities close to the mass-transport-limited current density, as we shall see below. In this way, one can create rough surfaces for which the resonance frequency shifts by as much as 30–40 kHz upon immersion in a dilute aqueous solution.

16.1.5 Scales of Roughness

In this context, it is important to consider the various scales of roughness, and the place of the EQCM on this scale. Starting from the finest scale, we consider the Helmholtz double-layer capacitance, adsorption of small molecules, and charge-transfer kinetics. All these phenomena occur within 1 nm from the surface, and are therefore sensitive to roughness on the atomic scale. Next, there is the Gouy–Chapman diffuse double layer, which can extend up to about 10 nm in a dilute (1 mM) solution of a 1 : 1 electrolyte, and essentially disappears in a 1 M solution. The Nernst diffusion layer is much thicker, in the range 5–100 μm. This is the reason that an electrode may have a high roughness factor, as far as adsorption or formation of a UPD layer is concerned, while the diffusion-limited current density can still be calculated correctly, ignoring this roughness and considering the apparent (i.e. the geometrical) surface area.

The characteristic length relevant to the vibration of the EQCM is related to the depth in solution where the liquid still oscillates as a result of the oscillation of the surface of the crystal. This has been calculated and is given by

$$\delta = \left(\frac{\eta}{\pi \rho f_0} \right)^{1/2} \tag{16.9}$$

For a dilute aqueous solution at room temperature this yields a value of 0.23 μm, placing it between the atomic scale and the scale relevant to mass-transport-controlled processes.

The vibrations of the AT-cut quartz crystal are in the plane parallel to the surface in contact with the liquid. The velocity vector of the liquid decreases with increasing distance from the surface and is given by

$$v(z) = v_0 \exp\left(-\frac{z}{\delta}\right) \tag{16.10}$$

Thus, it can be said that within a distance of about 1 μm these vibrations essentially die out. The roughness of the surface on the atomic scale is relatively easy to determine, by measurement of the maximum amount of some small molecules or atoms (e.g. in the UPD region, oxide formation and adsorption of CO or of atomic hydrogen). Unfortunately, the results of such measurements are not necessarily relevant to roughness in the context of the EQCM, because of the large difference in the characteristic length for adsorption (1 nm) and for the EQCM (0.23 μm).

16.2 Impedance Analysis of the EQCM

16.2.1 The Extended Equation for the Frequency Shift

An extended equation for the shift in frequency of an EQCM may be written as

$$\Delta f = \Delta f_{\mathrm{m}} + \Delta f_{\eta} + \Delta f_{\mathrm{p}} + \Delta f_{\mathrm{T}} + \Delta f_{\mathrm{R}} + \Delta f_{\mathrm{sl}} \tag{16.11}$$

We have already discussed the first four terms on the right-hand side of this equation. It should be obvious that the EQCM operates as a true "microbalance" only when all terms other than Δf_{m} are zero. In the context of electrochemical measurements it is relatively easy to maintain constant temperature and pressure, but the term Δf_{η} may be difficult to determine, because the density and viscosity of a layer of the liquid adjacent to the surface may change as a result of the electrochemical process occurring at the interphase, even if the viscosity and density in the bulk of the solution are maintained constant.

The effect of roughness, Δf_{R}, is also hard to assess because, as pointed out above, in practice a surface is never ideally flat, while in Eq. (16.7) Δf_{η} was derived for an ideally flat surface. Theories have been published regarding the effect of different models of roughness, but will not be discussed here. At any rate, this is of importance only when the surface roughness is known to change, or when it is changed on purpose to test the theory of roughness. Otherwise, the roughness could be maintained constant during the electrochemical process, so that $\Delta f_{\mathrm{R}} \approx 0$.

The last term in Eq. (16.11) represents the so-called "slippage" at the interphase between the surface of the EQCM and the electrolyte. This relates to yet another tacit assumption in Eq. (16.7), which was derived (as is often done in hydrodynamics) assuming the so-called *sticking boundary condition* to solve the differential equations of hydrodynamic flow. This implies that the first layer of liquid in contact with the vibrating crystal at the interphase moves at the same

velocity as the surface itself, in the direction parallel to the interphase. Further layers of solution move more slowly, following Eq. (16.10). It should be obvious that slippage leads to a smaller interaction of the vibrating surface with the liquid adjacent to it, which might be interpreted as a virtual decrease of the viscosity term, Δf_η, which leads to an increase of the resonance frequency. Indeed, one is often surprised to find that adsorption of certain organic molecules leads to a *positive* value of Δf, implying that Δm is actually negative, consistent with the notion of increased slippage, caused by replacement of the hydrophilic surface of the metal by a hydrophobic surface of the organic molecules adsorbed on it.

16.2.2 Other Factors Influencing the Frequency Shift

Although Eq. (16.11) includes six different factors that could affect the resonance frequency, there are additional phenomena that we might denote, somewhat naughtily, as Δf_{etc}. One of these factors is internal stress, created by absorption of hydrogen in a layer of palladium coated on one side of the EQCM. It is well known that absorbed atomic hydrogen in Pd causes an expansion of the lattice. If only one side of the crystal is coated by Pd (the other side being a gold surface), the crystal will bend when hydrogen is absorbed, creating an internal stress, which shows up as a large decrease of Δf, over and above that calculated for the amount of hydrogen absorbed. If both sides are coated by Pd, the effect is much smaller and may vanish if the crystal is planar on both sides (which is not always the case). A similar effect was observed when a thin layer of copper was plated on one side of the crystal, probably caused by the stress or strain in the coated film. This may be a general phenomenon, observed when the crystal parameters of the coated metal differ from those of the gold surface, which is deposited on the quartz crystal.

If the surface of the sample is porous, small amounts of water may be occluded in the pores, influencing the observed frequency shift. Nonuniform current distribution during metal deposition may lead to nonuniform thickness of the plated metal. These and some fine points in the design and operation of an electrochemical cell, where one side of the EQCM acts also as the working electrode, could all be lumped into the Δf_{etc} term, to interfere with the simple interpretation of the change of frequency observed as a change of mass.

16.3 Uses of the EQCM as a Microsensor

16.3.1 Advantages and Limitations

When the EQCM was first introduced, the only parameter measured was the frequency of the resonance. The shift of the frequency was interpreted, often erroneously, as the change of mass calculated from Eq. (16.2). Since then, however, the EQCM has been studied with a frequency response analyzer, which yields the impedance spectrum of the system. It turns out that the impedance is a complex number, so that the real and the imaginary parts can be determined separately. An equivalent way of treating the data is to plot the real part of the

admittance as a function of frequency. In this way, the peak represents the resonance frequency, f, while the width-at-half-height, Γ, represents the imaginary part of the admittance. Combining Eqs. (16.7) and (16.8) we find

$$\Delta f_\eta = -f_0^{3/2} \frac{(\rho_{\mathrm{fl}}\eta_{\mathrm{fl}})^{1/2}}{(\pi\rho_{\mathrm{q}}\mu_{\mathrm{q}})^{1/2}} \tag{16.12}$$

where the subscripts "fl" and "q" refer to the fluid and the quartz crystal, respectively. The corresponding width of the resonance is given by the similar expression

$$\Gamma = 2f_0^{3/2} \frac{(\rho_{\mathrm{fl}}\eta_{\mathrm{fl}})^{1/2}}{(\pi\rho_{\mathrm{q}}\mu_{\mathrm{q}})^{1/2}} \tag{16.13}$$

It follows that the ratio of width-at-half-height to the shift of the resonance frequency, upon immersion of an EQCM in a fluid, is simply

$$\frac{\Gamma}{\Delta f_\eta} = -2 \tag{16.14}$$

Note that immersing the oscillating crystal in a liquid loads it, because energy is transferred to the solution adjacent to the surface of the vibrating crystal. This leads to a lowering of the resonance frequency, just as expected for mass loading. However, the width-at-half-height increases.

In Figure 16.2 we present a series of admittance spectra, showing the effect of viscosity and density on the shape of the admittance curves. Different solutions were used, as listed in the caption to this figure. No potential was applied and no electrochemical reaction took place. Adsorption of the solvent or the solute may have taken place, but this would not account for any increase in the width, and to only a small shift of frequency, of the order of a few Hertz, while the total shift of the resonance frequency shown is in the range of several kilohertz.

The results of a different experiment, intended to show the effect of roughness, are shown in Figure 16.3. In this case, electroplating of gold on the gold surface of the EQCM was conducted at two different current densities. Figure 16.3a shows the admittance curves measured at different times during electroplating at a low current density of 20 μA cm^{-2}. The curves are seen to be shifted to lower frequencies with increasing time of plating, indicating an increase in mass deposited, but their shape is essentially unchanged.

In Figure 16.3b, electroplating of gold was conducted from the same solution, but at a much higher current density of 500 μA cm^{-2}, in order to create a rough surface. Here, the resonance frequency is also seen to shift to lower values, but the width of the resonance increases significantly, showing that this cannot be interpreted as the result of added mass by itself.

The total shift of frequency shown here can be written as

$$\Delta f = \Delta f_{\mathrm{m}} + \Delta f_{\mathrm{R}} \tag{16.15}$$

but we cannot separate the two terms on the right-hand side of this equation, unless we determine one of them independently.

It is important to note that determination of the width of the resonance can be critical for the use of the EQCM as a true microbalance. Relating Δf to the change

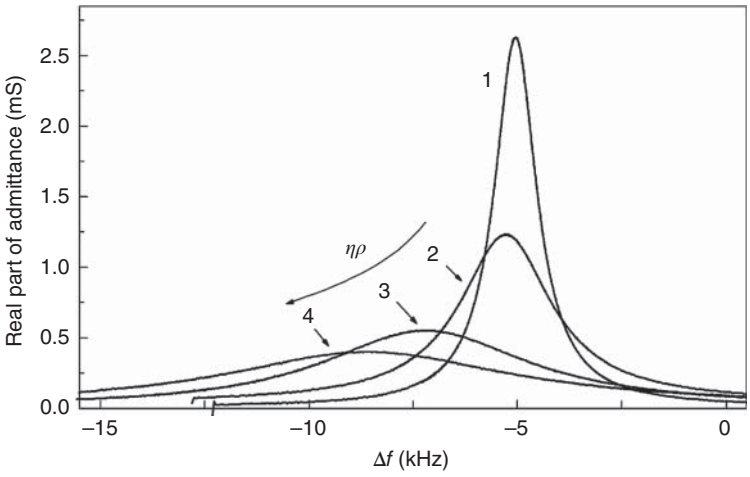


Figure 16.2 The effect of increasing viscosity of the fluid on the admittance curves. (1) Dimethyl ether, (2) water, (3) 40% aqueous solution of sucrose, (4) 50% aqueous solution of sucrose. Source: Reprinted with permission from L. Daikhin et al., *Anal. Chem*. **74** (2002) 554.

of mass Δm, employing Eq. (16.2), is only valid if the parameter Γ in Eq. (16.14) is constant, in a given composition of the solution and at constant temperature and pressure. In the general case, it may be better to regard the EQCM as a quartz crystal microsensor. On the other hand, measurement of both Δf and Γ can be very useful in more advanced analysis of the structure of the metal/electrolyte interphase, employing suitable models, which can be tested experimentally.

16.3.2 Some Applications of the EQCM

The high sensitivity and the real-time monitoring of mass changes on the sensor crystal make the QCM a very attractive technique for a large range of applications. Especially, the development of QCM systems for use in fluids or with viscoelastic deposits has dramatically increased the interest in this technique. One of the major advantages of the QCM technique for liquid systems is that it allows a label-free detection of molecules. There is a large range of applications of QCM and EQCM: (i) Thin-film thickness monitoring during thermal, e-beam, sputtering, magnetron, laser, and vacuum deposition, as well as plating and etching in liquid; (ii) biochemistry and biotechnology; (iii) drug delivery and drug research; (iv) in situ monitoring of lubricant and petroleum properties; (v) electrochemistry of interfacial processes at the metal/solution interphase; (vi) adhesion of cells, liposomes, and proteins; (vii) gas sensors; (viii) self-assembled monolayers; (ix) surfactant interactions with surfaces; (x) dissolution of polymer coatings.

The high sensitivity of the EQCM should make it an ideal tool for the study of trace amounts of impurities in air or in liquids. This is a large field by itself, and it is mentioned here only very briefly, to indicate some possible aspects of such applications. For a sensor to be useful, it should usually be selective to the specific impurity of interest. This can be achieved by depositing on the

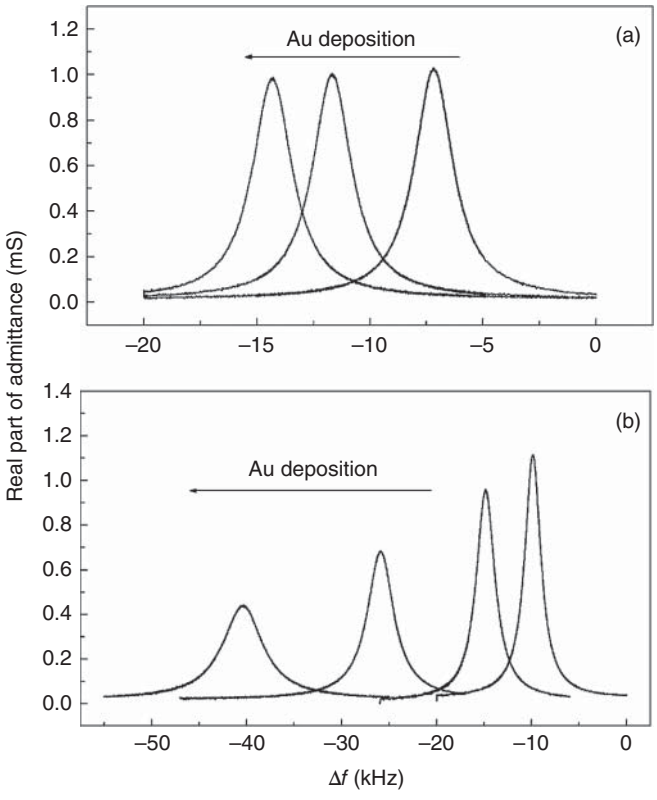


Figure 16.3 (a) Admittance spectra obtained at different times during plating of gold on the EQCM at a low current density of 20 μA cm^{-2}. (b) The same conducted at a higher current density of 500 μA cm^{-2}, which is close to the limiting current density, creating a rough surface. Source: Reprinted with permission from V. Tsionsky et al., Chapter 1 in *Electroanalytical Chemistry – A Series of Advances*, Vol. **22** (2004).

surface of the gold-coated quartz crystal a thin layer of a judiciously chosen material. A hygroscopic salt could be used to make it sensitive to humidity, a layer of palladium could make it sensitive to molecular hydrogen, a copper sulfate coating could detect ammonia in the gas, a high-surface-area platinized platinum coating could respond selectively to methanol, and so on. In biology and medicine, the use of antibodies and antigens could render the surface highly selective. The mass of a single bacterium is about 3 orders of magnitude below the detection limit of an EQCM, but the rate of growth of colonies of bacteria could be followed. In most such applications, the simple relationship between mass and frequency shift given by Eq. (16.2) may not apply, but the sensor could be calibrated to yield useful result.

Metal deposition may be expected to be most suitable to measure added mass by determining the shift in resonance frequency. The metal atoms electrodeposited become part of the crystal lattice of the substrate and are therefore considered to be rigidly attached to the surface. Thus, slippage is not an issue. The surface is generally not ideally flat, but as long as the width of the resonance

remains unchanged, it may be correct to use Eq. (16.2) to determine the added mass. The sensitivity is high enough to detect a fraction of a monolayer, and maintaining a current density well below the mass-transport-limited current density will usually keep the roughness unchanged.

In Figure 16.4 we show the results obtained for the deposition of copper, silver, and gold. In each case, the substrate is the same as the metal being deposited. If the EQCM acts in this system as a microbalance (i.e. if Eq. (16.2) is applicable), the coordinates of Δf can be recalculated to represent the mass increase, Δm. The slope of this line is the equivalent weight of the metal deposited. Any systematic deviation from the calculated line can be associated with a lower current efficiency, or some other deviation of the system from the conditions under which Eq. (16.2) is applicable.

The data for silver deposition agree with the calculated line within $\pm 0.25\%$. For deposition of copper, the data also follow a straight line, but the slope in this and similar experiments was found to show consistently a lower mass than the calculated slope, by about 6–8%. This deviation from theory was not attributed to the properties of the EQCM. The electroplating of copper is commonly believed to occur in two consecutive charge-transfer steps:

$$Cu^{2+}_{soln} + e^{-}_{M} \rightarrow Cu^{+}_{soln} \tag{16.16}$$

followed by

$$Cu^{+}_{soln} + e^{-}_{M} \rightarrow Cu^{0}_{M} \tag{16.17}$$

Most of the monovalent copper formed at the surface will be reduced to metallic copper, but some may diffuse away into the bulk of the solution, without being

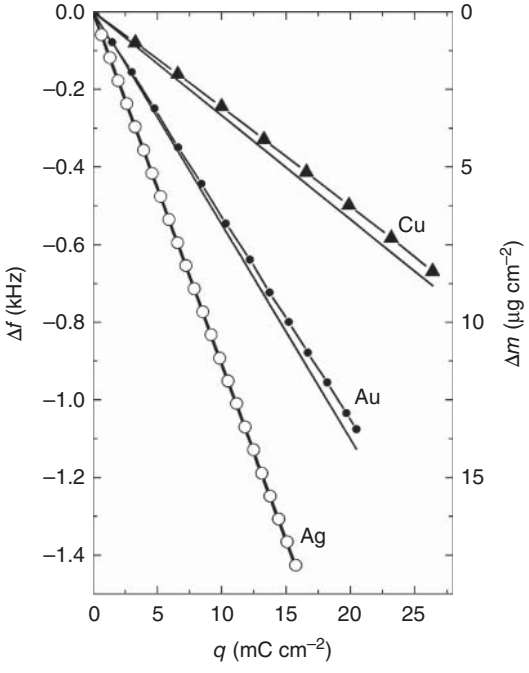


Figure 16.4 Plots of the frequency shift versus charge, during galvanostatic plating of Cu, Ag, and Au (each metal on its own metal substrate). Lines are calculated. Points represent experimental results. Source: Reprinted with permission from V. Tsionsky et al., Chapter 1 in *Electroanalytical Chemistry – A Series of Advances*, Vol. **22** (2004).

further reduced. Thus, a fraction of the electrons consumed in Eq. (16.16) may not be used to form atomic copper on the surface, and the effective charge will be less than that measured, leading to the lower slope. The same argument applies to deposition of Au, where the reactant is Au^{3+} and the intermediate contains Au^{+}.

EQCM measurements provide a very sensitive means for the determination of equivalent weight (EW), using the following equation:

$$\mathrm{EW} \equiv \frac{M}{n} = \frac{F\Delta m}{\Delta Q} \tag{16.18}$$

where M is the molar mass (g mol^{-1}), n is the number of electrons transferred per molecule (dimensionless), Δm is as defined for Eq. (16.2) (g cm^{-2}), and ΔQ is the charge density (C cm^{-2}). Thus, the EW can be deduced from the slope of the line Δm versus ΔQ.

Further Reading

1 Eliaz, N., Kopelovitch, W., Burstein, L. et al. (2009). Electrochemical processes of nucleation and growth of calcium phosphate on titanium supported by real-time quartz crystal microbalance measurements and X-ray photoelectron spectroscopy analysis. *J. Biomed. Mater. Res. A* 89: 270–280.

2 Urbakh, M., Tsionsky, V., Gileadi, E., and Daikhin, L. (2007). Probing the solid/liquid interface with the quartz-crystal microbalance. In: *Piezoelectric Sensors* (ed. C. Steinem and A. Jonshoff), 111–149. Berlin: Springer.

3 Olsson, C.-O.A. and Landolt, D. (2006). Electrochemical quartz crystal microbalance, Chapter 19. In: *Analytical Methods in Corrosion Science and Engineering* (ed. P. Marcus and F. Mansfeld), 733–751. Boca Raton, FL, USA: CRC Press.

4 Buck, R.P., Lindner, E., Kutner, W., and Inzelt, G. (2004). Piezoelectric chemical sensors (IUPAC Technical Report). *Pure Appl. Chem.* 76: 1139–1160.

5 Hillman, A.R. (2003). The electrochemical quartz crystal microbalance. In: *Encyclopedia of Electrochemistry*, Instrumentation and Electroanalytical Chemistry, vol. 3 (ed. A.J. Bard, M. Stratmann and P.R. Unwin), 230–289. Weinheim: Wiley-VCH.

6 Marx, K.A. (2003). Quartz crystal microbalance: a useful tool for studying thin polymer films and complex biomolecular systems at the solution-surface interface. *Biomacromolecules* 4: 1099–1120.

7 Daikhin, L., Gileadi, E., Tsionsky, V. et al. (2000). Slippage at adsorbate–electrolyte interface. Response of electrochemical quartz crystal microbalance to adsorption. *Electrochim. Acta* 45: 3615–3621.

8 Ward, M.D. (1995). Principles and applications of the electrochemical quartz crystal microbalance. In: *Physical Electrochemistry: Principles, Methods and Applications*, vol. 7 (ed. I. Rubinstein), 293–337. New York: Marcel Dekker.

9 Schumacher, R. (1990). The quartz microbalance: A novel approach to the in-situ investigation of interfacial phenomena at the solid/liquid junction. *Angew. Chem. Int. Ed.* 29: 329–343.

10 Kanazawa, K.K. and Gordon, G. (1985). The oscillation frequency of a quartz resonator in contact with a liquid. *Anal. Chim. Acta* 175: 99–105.
11 Sauerbrey, G. (1959). Use of vibrating quartz for thin film weighing and microweighing. *Z. Phys.* 155: 206–222. (in German).

17

Corrosion

17.1 The Definition of Corrosion

Corrosion, in its broad definition, is *the deterioration of a material due to a reaction with its environment*. This definition is so broad that phenomena such as hot corrosion and other high-temperature oxidation processes, cracking of polymers due to sunlight, polymers degradation in the low Earth orbit (LEO) environment due to exposure to atomic oxygen (AO), environmental stress cracking (ESC), and microbiological corrosion (MIC) of polymers, glass disease, attack of silicon dioxide by hydrofluoric acid, corrosion of composites, attack of a solid metal by another molten metal (liquid metal embrittlement), etc. can all be regarded as types of corrosion. However, here we shall limit ourselves to *corrosion of metals and alloys in aqueous solutions*, which is normally an *electrochemical process*.

Corrosion is a common phenomenon, observed all around us. It is found in pipes, automobiles, oilfield equipment, nuclear reactors, metallic implants, etc. Wherever there is a metal there is bound to be, sooner or later, corrosion. This is hardly surprising, since all metals, except gold, are thermodynamically unstable with respect to their oxides in air and in water. This is manifested by the observation that metals are not found in nature in their "native" or metallic form, but rather in the form of some compound, an oxide, a sulfide, a silicate, and so on.[1] The history of humankind is closely linked with the technology of reducing ores to the corresponding metals or alloys. This requires the input of energy, and the resulting product is unstable thermodynamically. Corrosion can be regarded as the natural tendency of metals to revert to a more stable state as a chemical compound of one kind or another, depending on the environment. Our technology, therefore, depends on our ability to slow down the rate of corrosion to an acceptable level.

The first and most fundamental step in corrosion is the oxidation of the metal to its lowest stable valence state. This is most often followed by the formation of insoluble products, the exact nature of which depends on the metal and on the environment in which it is corroding. The ion formed in the initial step may be oxidized further, producing oxides or other compounds of mixed valence. In some cases (e.g. Al, Ti, Cr) the corrosion products form dense insulating layers,

1 There are some exceptions, when base metals are found in a reducing environment, e.g. at the bottom of the oceans, but this is rather rare.

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

which prevent further corrosion. In other cases (e.g. low-carbon steel), the layer is porous, allowing the corrosion process to continue until the whole piece of metal has been consumed. When a protective layer does exist, it does not have to be very thick: about 5 nm in the case of Al and 3 nm in the case of stainless steel. Thus, as little as 10–20 molecular layers of the oxide film can provide excellent protection for long periods of time.

Anodic dissolution of a metal cannot occur *by itself* for any length of time, since it would lead to charging of the metal to a high negative potential.[2] The accompanying reaction in aqueous solutions is usually hydrogen evolution or oxygen reduction.

When a piece of iron is placed in 1.0 M HCl it dissolves readily, with simultaneous evolution of hydrogen. The rate of dissolution can be measured by determining the mass loss of iron, or by analysis of the solution for its ions, but one could also determine this quantity by measuring the volume of hydrogen evolved. The rates of anodic metal dissolution and of cathodic hydrogen evolution must be equal, because there can be no accumulation of electrons in the metal or the solution phase.

Let us discuss the corrosion process in more detail. Metal atoms characteristically lose electrons in what is known as *oxidation reaction*. For example, the hypothetical metal M that has a valence of n may experience oxidation according to the reaction

$$M \rightarrow M^{n+} + ne^- \tag{17.1}$$

in which M becomes an $n+$ positively charged ion or, cation, and in the process loses its n-valence electrons; e^- is used to symbolize an electron. Examples in which metals oxidize include

$$Fe \rightarrow Fe^{2+} + 2e^- \tag{17.2}$$

$$Al \rightarrow Al^{3+} + 3e^- \tag{17.3}$$

The site at which *oxidation* takes place is called the *anode*; accordingly, oxidation is sometimes called an *anodic reaction*.

The electrons generated from each metal atom that is oxidized must be transferred to and become a part of another chemical species, in what is termed a *reduction reaction*. For example, some metals undergo corrosion in acid solutions, which have a high concentration of hydrogen ions (H^+). The H^+ ions are reduced as follows:

$$2H^+ + 2e^- \rightarrow H_2 \tag{17.4}$$

and hydrogen gas, H_2, is evolved. Note that the hydrogen evolution reaction (HER) is written here in a simplified form. See Section 7.2 for more details.

Other reduction reactions are possible, depending on the nature of the solution to which the metal is exposed. For an acid solution having dissolved oxygen,

---

2 If one assumes a double-layer capacitance of 20 μF cm^{-2}, it would take 20 μC cm^{-2} to change the potential across the interface by 1.0 V. This amount of charge corresponds to 1×10^{-10} mol cm^{-2}, or 6 ng cm^{-2} of iron.

reduction according to

$$O_2 + 4H^+ + 4e^- \rightarrow 2H_2O \tag{17.5}$$

will probably occur. Or, for a neutral or basic aqueous solution in which oxygen is also dissolved,

$$O_2 + 2H_2O + 4e^- \rightarrow 4(OH)^- \tag{17.6}$$

Any metal ions present in the solution may also be reduced. For ions that can exist in more than one valence state (i.e. multivalent ions), reduction may occur by

$$M^{n+} + e^- \rightarrow M^{(n-1)+} \tag{17.7}$$

in which the metal ion decreases its valence state by accepting an electron. Alternatively, a metal may be totally reduced from an ionic to a neutral metallic state according to

$$M^{n+} + ne^- \rightarrow M \tag{17.8}$$

Note that at equilibrium, the rate of the anodic reaction must be equal to the rate of the cathodic reaction, because charge cannot be accumulated, either in the metal or in the solution.

The corrosion of iron and formation of rust ($F_2O_3 \cdot H_2O$) take place in several stages (see Figure 17.1). The first and most fundamental step is the oxidation of iron to its lowest stable valence state (see Eq. (17.2)). The Fe^{2+} cation diffuses away from the electrode surface, encounters dissolved O_2, and is further oxidized to Fe^{3+}. The latter is deposited as Fe_2O_3, following the reaction

$$2Fe^{3+} + 4H_2O \rightarrow Fe_2O_3 \cdot H_2O\,(s) + 6H^+\,(aq) \tag{17.9}$$

In alkaline (basic) environments, hydroxides may form too:

$$Fe^{2+} + 2(OH)^- \rightarrow Fe(OH)_2 \tag{17.10}$$

$$4Fe(OH)_2 + O_2 \rightarrow 2(Fe_2O_3 \cdot H_2O) + 2H_2O \tag{17.11}$$

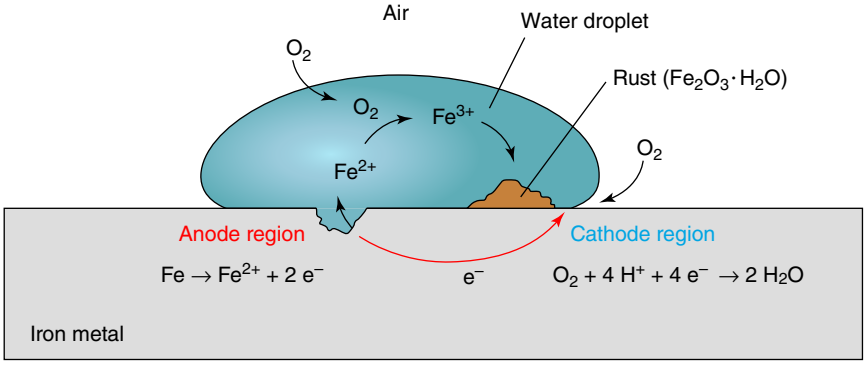


Figure 17.1 The mechanism of rusting.

Corrosion thus requires water and oxygen and is accelerated in the presence of salts. Note that the chemical composition of the metal oxide is not uniform across its thickness; it varies from the lowest valence of iron (FeO) at the interface with the metal, to Fe_3O_4, to Fe_2O_3 at the outer layer. The Fe_2O_3 layer formed as a result of oxidation is porous and flakes off, so it does not protect the metal beneath it from further oxidation.

17.2 Corrosion Costs

The cost to society of corrosion and its prevention is staggering,[3] as can be realized from the following published data. In England, every 90 seconds a ton of steel is converted into rust. On average, Israel loses 15% of its water supply due to corrosion. Corrosion cost analyses have been carried out around the world, and the findings have been similar – corrosion costs of 1.5–3.5% of the gross national product (GNP), 23–25% of which were claimed avoidable, were reported. A US study conducted by Battelle Columbus Laboratories and published in 1978 found that the total corrosion cost was 4.9% of the GNP of the United States, 40% of which was estimated to be avoidable (i.e. costs that can be reduced "by the most economically effective use of presently available corrosion technology").

A two-year study entitled "Cost of Corrosion and Preventative Strategies in the US" was conducted in 1999–2001 by CC Technologies, Inc. (Dublin, OH) per mandate of the US Congress.[4] The direct costs of metallic corrosion in several US industry sectors were first determined; these are costs incurred by owners and operators of structures, manufacturers of products, and suppliers of services. The problem is that indirect costs are not taken into account. These include cost of labor attributed to corrosion management activities, cost of the equipment required because of corrosion-related activities, loss of revenue due to disruption in supply of product, cost of loss of reliability, taxes and overhead on the cost of corrosion portion of goods and services, etc. Taking indirect costs into account approximately doubles the corrosion costs. This study found that the total direct annual cost of corrosion in the analyzed sectors was $137.9 billion (1.57% of the gross domestic product, GDP). This cost increased to $275.7 billion (3.14% of the GDP) when extrapolating to all national sectors. To put things in perspective, the annual costs of all the natural disasters combined (hurricanes, tornadoes, tropical storms, floods, fires, etc.) in the United States were $17 billion, i.e. 15 times lower than the corrosion costs! The per capita direct cost of corrosion for US residents was approximately $970 per year. It was also concluded that 25–30% of the annual corrosion costs in the United States could be saved if optimum corrosion management practices (namely, proper design, materials selection, and the use of existing corrosion control methodologies)

3 One should watch: Spotlight on: Corrosion, NACE International, https://www.youtube.com/watch?v=bKvroCboUVc and NACE International 2009, https://www.youtube.com/watch?v=5reVkdHIHs8.

4 See Federal Highway Administration (FHWA) Report FHWA-RD-01-156, September 2001, www.corrosioncost.com, and *Materials Performance*, Supplement, July 2002, for more details.

were employed. The problem cannot be eliminated, however, since as pointed out earlier, corrosion represents the natural tendency of all systems toward a state of minimum Gibbs energy. Therefore, one should use terms such as corrosion control or corrosion management, and not corrosion prevention.

The damage caused by corrosion is of two general kinds: aesthetic and engineering. An example of the former is the development of rust spots on (so-called) stainless steel cutlery. Although rust, which is just a mixture of oxides of iron, is not harmful, one would not like to eat with a rusted fork or spoon. Examples of engineering damage due to corrosion are countless. From car bodies to pipelines to electronic components, almost everything must be protected and eventually replaced because of corrosion.

In most cases, corrosion can be effectively eliminated by investment in the construction material, or at least slowed down to the point that the device, be it a piece of machinery or a structure, will have to be replaced for some other reason, before corrosion has become severe. Jewelry and coinage are extreme examples, but even in a chemical plant one has the choice of designing for minimum maintenance and long periods between overhauls at a high initial cost, or frequent maintenance at a lower initial investment.

There are "corrosive" environments and those that are considered benign. The combination of high humidity and high temperature favors corrosion, but above all, the presence of chloride ions (and other halides) is detrimental to almost all metals and interferes with many methods of corrosion protection. Chloride is not the only ion that enhances corrosion but it is the one most commonly found all around us, in seawater and even in freshwater, in the ground, and in the human body. Salt spray carried by the wind from the sea is a major cause of corrosion, and it is easy to see how the importance of this factor diminishes with increasing distance inland.

The rate of *uniform corrosion* (also known as general corrosion) can be expressed in terms of thickness reduction per time unit.[5] In the case of uniform corrosion of construction steel, corrosion rate lower than 5 mpy is usually acceptable, whereas corrosion rate higher than 50 mpy represents a significant problem. However, corrosion is often not uniformly distributed on the exposed surface (i.e. it is *localized corrosion*). In this case, an average corrosion rate of 2 mpy might be concentrated in spots, leading to holes in a piece of metal (e.g. a pipeline) that is 200 mil (about 0.5 cm) thick or more. A particular type of local corrosion occurs when the worst design error, from the corrosion point of view, is made – bringing in contact two different metals without isolating them from each other electrically. For example, connecting copper plates with steel rivets in effect creates a battery in which steel is the anode and copper is the cathode. Moreover, since the two metals are in intimate contact with each other, this is equivalent to having the terminals of this battery effectively shorted, leading to a high rate of discharge (i.e. corrosion). As long as the structure is totally dry, nothing will happen; but, if water accumulates on the surface, corrosion of the rivets might occur, leading to critical structural damage.

---

5 The corrosion rate (CR) is commonly expressed in units of mpy (mil per year). Note that 1 mil = 0.001 in. = 25.4 μm.

17.3 Thermodynamics of Corrosion

17.3.1 Introduction and Important Terms

In this section we will focus on the thermodynamics of electrode reactions. All potentials will be given against the standard hydrogen electrode (SHE), and the concentrations will be written in square brackets, for clarity. The potential of phases, including the meaning of the SHE scale (see Section 2.3), was already discussed in Chapter 2. Therefore, this topic is discussed only briefly in this section. Thermodynamics can provide only the negative answers; it allows us to calculate and determine which reactions *will not* happen. We need kinetic information to determine what *will* happen at a rate that may be of practical interest, or at least at a rate that can be detected. Therefore, complementary discussion on the kinetics of corrosion will be presented in Section 17.4.

Let us begin with a quick recap on electrolytes, acid, and bases. Electrolytes are molecules that ionize in water. For example, a salt such as NaCl would produce an electrolyte solution. Those compounds that produce a large number of ions in solution are called strong electrolytes. On the other hand, those compounds that produce a small number of ions in solution are weak electrolytes. The concentration of electrolyte solutions can be expressed in various ways: (i) weight per volume (e.g. 8.5 g L^{-1} NaCl); (ii) percentages, either weight or volume, of solute in a given solution (e.g. 5% w/v means 5 g of solute to which solution is added to 100 mL); (iii) molarity, which measures the number of moles of solute per 1 L of solution (mixture). For example, the molecular weight (MW) of glucose is 180.2 g mol^{-1}, i.e. 1 mol of glucose weighs 180.2 g, and 1 M solution of glucose is prepared by weighing 180.2 g glucose and adding to it DI water, bringing the mixture to 1 L. (iv) Equivalents: 1 Eq. will neutralize 1 mol of H^+ or $(OH)^-$ ions. This unit depends on both the concentration and electric charge of the ion, by multiplying the molar concentration by the valence of the ion. For example, 1 mM Na^+ = 1 mEq L^{-1}, 1 mM Ca^{2+} = 2 mEq L^{-1}. (v) Molality (m): expresses the number of moles of solute per kilogram of solvent. For example, 0.5 m KNO_3 means that 0.1 mol of KNO_3 are dissolved in 200 g of water. The primary advantage of using molality as a measure of concentration is that molality only depends on the masses of solute and solvent, which are unaffected by variations in temperature and pressure (in contrast to solutions prepared volumetrically). Other advantages of molality are that the molality of one solute in a solution is independent of the presence or absence of other solutes, and that chemical reactions take place in proportion of mass, not volume.

In Chapter 2 we defined the relation between activity and concentration:[6]

$$a_i = \gamma_i c_i \tag{2.3}$$

The *ionic strength* of a solution, I, is a measure of the amount of ions present (namely, the total concentration of charge in the solution). It is defined as

$$I = \frac{1}{2} \sum_i m_i z_i^2 \tag{17.12}$$

6 Unless otherwise stated, concentrations and partial pressures, instead of activities and fugacities, respectively, are used here and in all following equations, for simplicity.

where m_i is the molality of ion i and z_i is the charge on ion i. The ionic strength emphasizes the charges of ions because the charge numbers occur as their squares.

The chemical potential or activity of ions cannot be determined on a purely thermodynamic basis because the effects of an ion cannot be separated from the effects of the accompanying counterion (namely, solutions must be electrically neutral). While the individual activity coefficients of the individual ionic species cannot be measured experimentally, one can either measure or calculate the *mean activity coefficient* $\gamma_{\pm}$, which is the geometric mean (the square root of the product) of the activity coefficients of the individual ionic species. Let us demonstrate it for NaCl. The chemical potentials of its species in solution follow the relation

$$\mu_{\mathrm{NaCl(aq)}} = \mu_{\mathrm{Na^+(aq)}} + \mu_{\mathrm{Cl^-(aq)}} \tag{17.13}$$

The mean activity coefficient in this case is

$$\gamma_{\pm} = \sqrt{\gamma_+ \cdot \gamma_-} \tag{17.14}$$

The activity of such an electrolyte may thus be defined as

$$a_{\mathrm{salt}} = m_+ \cdot m_- \cdot \gamma_{\pm}^2 \tag{17.15}$$

But, for a NaCl solution of molality m we have $m_+ = m_- = m$, so that

$$a_{\mathrm{salt}} = m^2 \cdot \gamma_{\pm}^2 \tag{17.16}$$

Similarly, for a $ZnCl_2$ electrolyte one would get

$$a_{\mathrm{salt}} = m_+ m_-^2 \cdot \gamma_-^2 \gamma_+ \tag{17.17}$$

and

$$\gamma_{\pm} = \sqrt[3]{\gamma_+ \cdot \gamma_-^2} \tag{17.18}$$

Thus

$$a_{\mathrm{salt}} = m \cdot (2m)^2 \cdot \gamma_{\pm}^3 = 4m^3 \gamma_{\pm}^3 \tag{17.19}$$

Similarly, for an A_xB_y electrolyte that dissolves in an aqueous solution as $A_xB_y \rightarrow xA^{m+} + yB^{n-}$ we obtain

$$\gamma_{\pm} = (\gamma_+^p \cdot \gamma_-^q)^{1/s}, \quad \text{where } s = (p+q) \tag{17.20}$$

Values of the mean activity coefficients are tabulated in numerous publications. The behavior of an electrolyte solution deviates considerably from that of an ideal solution as a result of electrostatic forces. In 1923, Peter Debye and Erich Hückel developed a theory that allows us to calculate the mean activity coefficient of the solution. The Debye–Hückel theory is based on several assumptions of how ions act in solution: (i) departure from ideality is due to electrostatic interactions between ions, mediated by Coulomb's law. (ii) Strong electrolytes are assumed to be completely dissociated into ions. (iii) Ions are spherical and are not polarized by the surrounding electric field. Solvation of ions is ignored except insofar as it determines the effective sizes of the ions. (iv) Each cation is surrounded by a

spherically symmetric cloud of other ions with a net negative charge. Similarly, each anion is surrounded by a cloud with net positive charge. (v) The dielectric constant of the pure solvent is used. Changes caused by the presence of dissolved ions are not taken into account.

The mean activity coefficient can be calculated at very low solute concentrations ($I < 10^{-2.3}\ m$, ~0.005 m) by the Debye–Hückel limiting law:

$$\log_{10}\ \gamma_{\pm} = -A\,|z_+ \cdot z_-| \cdot \sqrt{I} \tag{17.21}$$

The constant A is calculated as

$$A = 1.825 \times 10^6 (\varepsilon_w T)^{-3/2} \tag{17.22}$$

where ε_w is the dielectric constant of water and T is the absolute temperature. For water at 298 K, $\varepsilon_w = 78.54$ and $A = 0.509\ \text{mol}^{-1/2}\ \text{kg}^{1/2}$.

At a higher ionic strength ($10^{-2.3}\ m < I < 0.1\ m$), the mean activity coefficient can be calculated by the Extended Debye–Hückel law:

$$\log_{10}\gamma_{\pm} = \frac{-A\,|z_+ \cdot z_-| \cdot \sqrt{I}}{1 + Ba_0\sqrt{I}} \tag{17.23}$$

For water at 25 °C, $B = 3.290 \times 10^9\ \text{mol}^{-1/2}\ \text{kg}^{1/2}\ \text{m}^{-1}$. In Eq. (17.23), a_0 is the average of the effective ion diameters of the cation and the anion (in meters). Large ions and less highly charged ions bind water less tightly and have smaller hydrated radii than smaller, more highly charged ions. Typical values are 3 Å for ions such as K^+, Cl^-, CN^-, and $HCOO^-$. As the product Ba_0 is not far from unity, Güntelberg introduced a simplified version of Eq. (17.23), i.e. without Ba_0 in the denominator. For even higher concentrations of electrolytes ($I \leq 0.5\ m$), the Davies law applies:

$$\log_{10}\gamma_{\pm} = -A\,|z_+ \cdot z_-| \cdot \left(\frac{\sqrt{I}}{1+\sqrt{I}} - 0.3I\right) \tag{17.24}$$

Acids are proton donors. Strong acids (e.g. HCl) ionize freely and markedly lower the pH, whereas weak acids ionize only slightly. Bases are proton acceptors. Strong bases (e.g. NaOH) ionize freely and markedly raise the pH, whereas weak bases ionize only slightly. In Chapter 6 we defined pH as

$$\text{pH} \equiv -\log a_{H^+} \tag{6.38}$$

The pH scale for aqueous solutions and natural waters is often given as 0–14. However, values of pH less than 0.0 and greater than 14.0 not only are possible but also have been prepared. For example, commercially available concentrated HCl solution (37% by mass) has $\text{pH} \approx -1.1$, while saturated NaOH solution has $\text{pH} \approx 15.0$.

The three most important equilibria for water are the following:

1) Self-ionization

$$H_2O \rightleftharpoons H^+ + OH^- \tag{17.25}$$

for which the equilibrium constant at 25 °C is given by

$$K_w = [H^+] \times [OH^-] = 1.008 \times 10^{-14} \text{ mol}^2 \text{ L}^{-2} \tag{17.26}$$

Neutrality is defined when $[H^+] = [OH^-]$. Thus, pH = 7 for neutrality at 25 °C. However, because the equilibrium constant (ionic product) is a function of temperature, the pH at neutrality depends on temperature too. For example, $K_w = 0.114 \times 10^{-14}$ $(\text{mol L}^{-1})^2$ at 0 °C and 5.476×10^{-14} $(\text{mol L}^{-1})^2$ at 50 °C; thus, the corresponding pH values at neutrality will be 7.47 and 6.63, respectively.

2) The HER, for which the Nernst equation is

$$E_{rev} = \frac{2.3RT}{2F} \log \frac{[H^+]^2}{[H_2]} \tag{17.27}$$

This equation can also be written in the form

$$E_{rev} = -\frac{2.3RT}{2F} \log[H_2] - \frac{2.3RT}{F}\text{pH} \tag{17.28}$$

3) The oxygen evolution reaction (OER), for which

$$E_{rev} = +1.229 + \frac{2.3RT}{4F} \log[O_2] - \frac{2.3RT}{F}\text{pH} \tag{17.29}$$

where $[H_2]$ and $[O_2]$ represent the partial pressures of hydrogen and oxygen, respectively.

Equations (17.30)–(17.32) demonstrate the meaning of the equilibrium constant for different types of reaction:

1) Dissociation in gaseous phase

$$2H_2O\,(g) \rightleftharpoons 2H_2\,(g) + O_2\,(g) \qquad K = \frac{P_{H_2}^2 \times P_{O_2}}{P_{H_2O}^2} \tag{17.30}$$

2) Ions in solution

$$H_2O\,(l) \rightleftharpoons H^+\,(aq) + OH^-\,(aq) \qquad K = c_{H^+} \times c_{OH^-} \tag{17.31}$$

3) Ions and solids

$$Fe(OH)_2(s) \rightleftharpoons Fe^{2+}(aq) + 2(OH)^-(aq) \qquad K = c_{Fe^{2+}} \times c_{OH^-}^2 \tag{17.32}$$

The following equations are useful in thermodynamic calculations of electrode reactions. For brevity, we include them in this section without showing how they were obtained:

$$\log K = -\frac{\sum \nu_\alpha \mu_\alpha^0 \,(\text{cal mol}^{-1})}{4.575\ T} \tag{17.33}$$

$$\log K = -\frac{\sum \nu_\alpha \mu_\alpha^0 \,(\text{cal mol}^{-1})}{1363} \quad \text{at } T = 25°\text{C} \tag{17.34}$$

$$E^0\,(\text{V}) = -\frac{\Delta\mu^0\,(\text{cal mol}^{-1})}{23{,}060\ n} \tag{17.35}$$

The fundamental thermodynamic function termed *chemical* potential, μ (J mol^{-1}), was defined for mixtures of different constituents. The energy of the system was differentiated with respect to the macroscopic mass of the substance, at constant entropy S and volume V. One meaning of the chemical potential is the change in a characteristic thermodynamic state function per change in the number of molecules. Depending on the experimental conditions, the characteristic thermodynamic state function is either internal energy U, Helmholtz free energy F, Gibbs free energy G, or enthalpy H. In his main paper Gibbs actually used the term "intrinsic potential," and not "chemical potential." By this, he referred to a derivative that is "entirely determined at any point in a mass by the nature and state of the mass about that point." For a multicomponent system with N components ($k = 1, \ldots, N$), the chemical potential μ_i of substance i may be expressed in terms of partial derivatives as

$$\mu_i = \left(\frac{\partial U}{\partial n_i}\right)_{V,S,n_{k\neq i}} = \left(\frac{\partial F}{\partial n_i}\right)_{T,V,n_{k\neq i}} = \left(\frac{\partial G}{\partial n_i}\right)_{T,p,n_{k\neq i}} \equiv \overline{G}_i \tag{17.36}$$

where n is the number of moles (mol), p is pressure (Pa), T is the thermodynamic (absolute) temperature (K), $n_{k\neq i}$ denotes that all ns except n_i are kept constant, and $\overline{G}_i$ is the partial Gibbs molar free energy of component i.

17.3.2 Electrode Potentials and the Standard Electromotive Force (EMF) Series

The *electrode potential* (or electrochemical potential) results from the electrical double layer (EDL) at the metal/solution interphase and indicates the tendency of a metal toward electrochemical reactions. When equilibrium condition is reached, the net current that crosses the double layer is zero. The potential of a reaction at equilibrium is called the *reversible electrode potential*. A reversible cell is a type of electrochemical cell (also called galvanic cell or voltaic cell) in which reversing the current reverses the cell reaction. This potential can be predicted theoretically by the Nernst equation.

The *electromotive force (EMF) series* is generated by arranging the standard potentials of different half-cell reactions at *standard conditions* ($[M^{n+}] = 1$ M, $p = 1$ atm, $T = 25$ °C) relative to the SHE. According to the International Union of Pure and Applied Chemistry (IUPAC), this series is expressed in the form of *reduction* potentials, from positive values (noble metals) at the top to negative values (active metals) at the bottom (see Table 17.1). The potential represents the tendency of the reaction to occur; it is an intensive thermodynamic property (thus, it should not be multiplied by the number of electrons transferred – it is on a "per electron" basis). E^0 represents the maximum electrical energy that can be produced from the cell. E^0 of the cell must be positive in order that the overall cell reaction will occur spontaneously in the direction written.

We can calculate the overall standard cell potential by selecting the relevant half-cell reactions from the EMF series, changing the sign of a half-cell reaction when it is an oxidation reaction in the electrochemical cell, multiplying the anodic and cathodic half-cell reactions by the stoichiometric coefficients required

Table 17.1 The standard EMF series (reduction potentials).

Half-cell reaction	Standard potential E^0 (V versus SHE)
$F_2(g) + 2e^- \rightleftharpoons 2F^-(aq)$	+2.870
$H_2O_2(l) + 2H^+(aq) + 2e^- \rightleftharpoons 2H_2O(l)$	+1.776
$Au^{3+}(aq) + 3e^- \rightleftharpoons Au(s)$	+1.498
$PbO_2(s) + 4H^+(aq) + 2e^- \rightleftharpoons Pb^{2+}(aq) + 2H_2O(l)$	+1.467
$Cl_2(g) + 2e^- \rightleftharpoons 2Cl^-(aq)$	+1.358
$O_2(g) + 4H^+(aq) + 4e^- \rightleftharpoons 2H_2O(l)$ pH = 0	+1.229
$Pt^{3+}(aq) + 3e^- \rightleftharpoons Pt(s)$	+1.200
$Ag^+(aq) + e^- \rightleftharpoons Ag(s)$	+0.799
$Hg_2^{2+}(aq) + 2e^- \rightleftharpoons 2Hg(l)$	+0.788
$Fe^{3+}(aq) + e^- \rightleftharpoons Fe^{2+}(aq)$	+0.771
$O_2(g) + 2H_2O(l) + 4e^- \rightleftharpoons 4OH^-(aq)$ pH = 14	+0.401
$Cu^{2+}(aq) + 2e^- \rightleftharpoons Cu(s)$	+0.337
$Sn^{4+}(aq) + 2e^- \rightleftharpoons Sn^{2+}(aq)$	+0.150
$2H^+(aq) + 2e^- \rightleftharpoons H_2(g)$	**0.000**
$Pb^{2+}(aq) + 2e^- \rightleftharpoons Pb(g)$	−0.126
$Sn^{2+}(aq) + 2e^- \rightleftharpoons Sn(g)$	−0.136
$Ni^{2+}(aq) + 2e^- \rightleftharpoons Ni(s)$	−0.250
$Co^{2+}(aq) + 2e^- \rightleftharpoons Co(s)$	−0.277
$Cd^{2+}(aq) + 2e^- \rightleftharpoons Cd(s)$	−0.403
$Fe^{2+}(aq) + 2e^- \rightleftharpoons Fe(s)$	−0.440
$Cr^{3+}(aq) + 3e^- \rightleftharpoons Cr(s)$	−0.744
$Zn^{2+}(aq) + 2e^- \rightleftharpoons Zn(s)$	−0.763
$2H_2O(l) + 2e^- \rightleftharpoons H_2(g) + 2OH^-(aq)$	−0.828
$Mn^{2+}(aq) + 2e^- \rightleftharpoons Mn(s)$	−1.180
$Ti^{2+}(aq) + 2e^- \rightleftharpoons Ti(s)$	−1.630
$Al^{3+}(aq) + 3e^- \rightleftharpoons Al(s)$	−1.662
$Mg^{2+}(aq) + 2e^- \rightleftharpoons Mg(s)$	−2.363
$Na^+(aq) + e^- \rightleftharpoons Na(s)$	−2.714
$K^+(aq) + e^- \rightleftharpoons K(s)$	−2.925
$Li^+(aq) + e^- \rightleftharpoons Li(s)$	−3.040

to ensure that each involves the same quantity of charge, and summing up the half-cell potentials. Note than when we multiply the half-cell reactions by the stoichiometric coefficients, the change in free energy and the number of coulombs both increase by the same factor, leaving the potential (voltage) unchanged. However, if we combine two half-reactions of the same type (i.e. reduction or oxidation) to obtain a third half-reaction, the E^0 values are not additive, since this third half-reaction is not accompanied by another half-reaction that causes charges to cancel. In such a case, since free energies are always additive, we can combine

them and use Eq. (17.38) below to find the cell potential. For example, in order to determine E^0 for the reduction reaction $Fe^{3+}(aq)/Fe(s)$, we can use the half-cell reactions $Fe^{3+}(aq)/Fe^{2+}(aq)$ ($E^0 = +0.771\ V$) and $Fe^{2+}(aq)/Fe(s)$ ($E^0 = -0.440\ V$) from Table 17.1. Simply adding the two E^0 values would give a wrong value of $E^0 = +0.331$ V for the third reaction. However, if we add $\Delta G_1^0 = -0.771F$ and $\Delta G_2^0 = +0.880F$, we get $\Delta G_3^0 = +0.109F$, and considering that three electrons are transferred in this reaction we get the correct value, $E^0 = -0.036$ V.

Let us consider the immersion of a pure copper electrode in an electrolyte solution of 1 M $CuSO_4$ at standard conditions. If we do not pass a current, we will measure $E_{rev} = E^0 = +0.337$ V on the voltmeter (see Table 17.1).

The meaning of overpotential was explained in Chapter 5, where it was defined by the following equation:

$$\eta \equiv E_{app} - E_{rev} \tag{5.2}$$

Overpotential represents the deviation from equilibrium. $\eta > 0$ would mean anodic polarization (namely, oxidation reactions will be favored), while $\eta < 0$ would mean cathodic polarization (namely, reduction reactions will be favored). For example, let us consider again the immersion of a pure copper electrode in an electrolyte solution of 1 M $CuSO_4$ at standard conditions. If $\eta < 0$, copper deposition from the solution will occur on the electrode, whereas if $\eta > 0$, the copper electrode will dissolve (corrode).

Next, let us consider a cell consisting of pure Zn and Cu electrodes, each immersed in an electrolyte solution containing 1 M of its cations (this is known as the Daniell cell; Figure 17.2). If the electrodes are connected electrically, Zn will undergo oxidation (corrode) and Cu will undergo reduction (electrodeposit; see Table 17.1). The overall reaction is

$$Zn(s) + Cu^{2+}(aq) \rightarrow Zn^{2+}(aq) + Cu(s) \tag{17.37}$$

When a current passes through the external circuit, electrons generated from the oxidation of zinc flow to the copper cell in order that Cu^{2+} be reduced. At

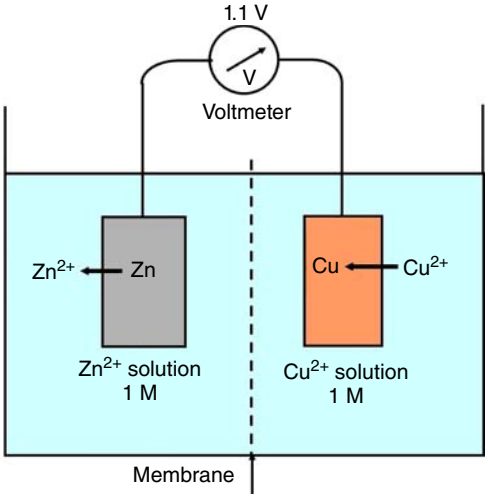


Figure 17.2 The Daniell electrochemical cell.

the same time, there will be some net motion of ions from each cell to the other across the porous membrane. This is called a *galvanic couple*, namely two metals electrically connected in an aqueous electrolyte wherein one metal becomes an anode and corrodes while the other acts as a cathode. An electric potential or voltage will exist between the two half-cells, and its magnitude can be determined if a voltmeter is connected in the external circuit. In the case of the Daniell cell, a standard potential of 1.1 V is expected based on the EMF series.

In order to make it easier to describe a given electrochemical cell, a special symbolic notation has been adopted, for example, $Zn|Zn^{2+}\|Cu^{2+}|Cu$. As a matter of convention, the oxidation reaction is written on the left side, and the reduction reaction on the right side. The vertical bars indicate phase boundaries, while the double vertical bar in the middle denotes the phase boundary between the two solutions.

An electrochemical cell such as that in Figure 17.2 has advantages over direct contact between the two electrodes: (i) part of the chemical energy can be converted directly to electrical energy, while in direct contact all of the released energy will be dissipated as heat; and (ii) deposits on the electrodes might terminate the reaction when the electrodes are in direct contact.

We should distinguish between galvanic cells and electrolytic cells. *Galvanic cells* use spontaneous reactions ($\Delta G < 0$) to generate electrical energy. The anode is the negative terminal, while the cathode is the positive terminal. A salt bridge usually acts to complete the circuit by joining the two half-cells together. Battery discharge and fuel cells are two examples of galvanic cells. In contrast, *electrolytic cells* use an applied voltage to drive a nonspontaneous ($\Delta G > 0$) chemical reaction. The electric current is supplied by an external source. In this case, the anode is the positive terminal, while the cathode is the negative terminal. Electroplating, water electrolysis, and battery charging are three examples of electrolytic cells. Both in the Pourbaix diagrams and in the E versus j curves, higher (more positive) potentials fortify oxidation reactions. Irrespective of whether the electrochemical cell is of the galvanic or the electrolytic type, the anode is always the electrode on which oxidation reactions (e.g. metal dissolution or oxygen evolution) take place, while the cathode is always the electrode on which reduction reactions (e.g. metal deposition or hydrogen evolution) take place.

The EMF series has several limitations: (i) it reflects standard conditions, while in situations of practical interest the system is rarely, if ever, under standard conditions; (ii) it lists only pure metals, not alloys; and (iii) it does not reflect kinetic effects such as passivity.

Let us elaborate on the latter argument. Figure 17.3a ranks 43 elements according to their thermodynamic and practical nobilities. From the practical standpoint, the nobility is higher as the immunity and passivation regions in the Pourbaix diagrams extend below and above the water stability region (see Section 17.3.5), and as these regions overlap with the pH range of 4–10, which is relevant to most applications. The practical effect of passivity on the following 10 metals is evident: Al, Be, Cr, Ga, Hf, In, Nb, Ta, Ti, and Zr. Yet, one should be careful when using Figure 17.3a because it is based on equilibrium diagrams, which are themselves only approximation, and because the corrosion and passivation reactions that are represented could be irreversible. Figure 17.3b

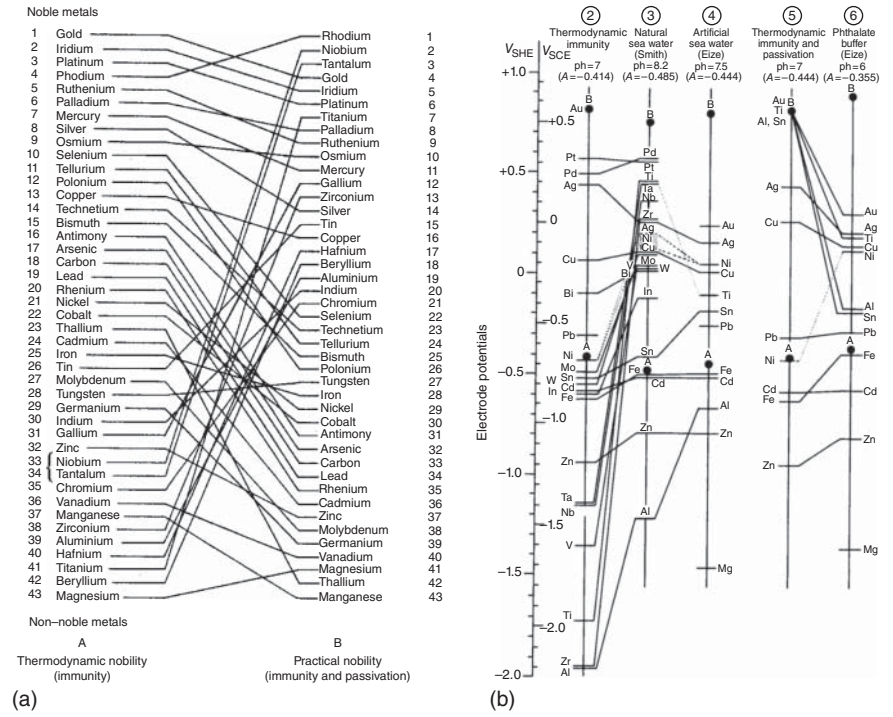


Figure 17.3 (a) Classification of metals and metalloids in order of thermodynamic nobility versus practical nobility. Source: Reprinted with permission from M. Pourbaix, *Atlas of Electrochemical Equilibria in Aqueous Solutions*, NACE, 1974. (b) Extension of (a) for different aqueous media and pH values. Source: Reprinted with permission from M. Pourbaix, *Lectures on Electrochemical Corrosion*, Plenum Press, 1973.

demonstrates the importance of comparing between the nobility of different metals in the specific environment of interest. From columns 2 and 5 it is evident that Cd is more noble than Fe at pH = 7. However, in natural seawater (pH = 8.2), artificial seawater (pH = 7.5), and phthalate buffer (pH = 6.0), Cd is more active (see columns 3, 4, and 6, respectively). In these three cases, Cd can be used as a sacrificial anode, protecting Fe from corrosion.

When two dissimilar metals in electrical contact are exposed to a common electrolyte, one of the metals can undergo increased corrosion while the other can show decreased corrosion. This type of accelerated corrosion is referred to as *galvanic corrosion*. The *galvanic series* is more appropriate than the EMF series for description of actual corrosion problems as alloys, and not pure metals, are often coupled, and the metals undergo corrosion rather than being in equilibrium with their ions at standard concentrations. ASTM G82-98(2014) is a common standard for construction and use of the galvanic series for predicting the galvanic corrosion performance. Usually, the further apart two metals are in the series (i.e. the greater the potential difference between them), the greater is the driving force for galvanic corrosion, and the greater the damage to the anode will often be. The anode-to-cathode surface areas ratio is also important: the lower this ratio is, the

higher the current density on the anode is, and the higher the corrosion rate is. The maximal damage will occur in the vicinity of the plane of contact between the two metals.

17.3.3 The Dependence of Free Energy on the Equilibrium Constant and Cell Potential

The change in the free energy of an electrochemical reaction can be determined either based on the overall cell potential or based on the equilibrium constant for that reaction:

$$\Delta G^0 = -nFE^0 = -RT \ln K \tag{17.38}$$

Note that $1\,\mathrm{V} = 1\,\mathrm{J\,C^{-1}}$ and that n is the number of electrons transferred in the reaction per mole of reactant. Thus, thermodynamic information can be deduced from electrochemical measurements, and vice versa.

17.3.4 The Nernst Equation

The *Nernst equation* allows us to calculate the cell potential for concentrations, partial pressures, temperatures, and pH values that are not standard. It is named after Walther Hermann Nernst (1864–1941) and can be easily derived as follows. Let us consider the reaction

$$a\mathrm{A} + b\mathrm{B} \rightleftharpoons c\mathrm{C} + d\mathrm{D} \tag{17.39}$$

The change in free energy is

$$\Delta G = G_{\text{products}} - G_{\text{reactants}} = (cG_{\mathrm{C}} + dG_{\mathrm{D}}) - (aG_{\mathrm{A}} + bG_{\mathrm{B}}) \tag{17.40}$$

For pure metals at standard conditions

$$\Delta G^0 = (cG^0_{\mathrm{C}} + dG^0_{\mathrm{D}}) - (aG^0_{\mathrm{A}} + bG^0_{\mathrm{B}}) \tag{17.41}$$

Subtracting Eq. (17.41) from Eq. (17.40) we get

$$\Delta G - \Delta G^0 = (cG_{\mathrm{C}} - cG^0_{\mathrm{C}}) + (dG_{\mathrm{D}} - dG^0_{\mathrm{D}}) - (aG_{\mathrm{A}} - aG^0_{\mathrm{A}}) - (bG_{\mathrm{B}} - bG^0_{\mathrm{B}}) \tag{17.42}$$

But, for example,

$$(cG_{\mathrm{C}} - cG^0_{\mathrm{C}}) = c\Delta G_{\mathrm{C}} = c(RT \ln a_{\mathrm{C}}) = RT \ln a^{\mathrm{c}}_{\mathrm{C}} \tag{17.43}$$

Thus, we can write for the cell reaction

$$\Delta G - \Delta G^0 = RT \ln \frac{a^{\mathrm{c}}_{\mathrm{C}} \cdot a^{\mathrm{d}}_{\mathrm{D}}}{a^{\mathrm{a}}_{\mathrm{A}} \cdot a^{\mathrm{b}}_{\mathrm{B}}} \tag{17.44}$$

Combining with Eq. (17.38) we get

$$-nFE_{\text{rev}} + nFE^0 = RT \ln \frac{a^{\mathrm{c}}_{\mathrm{C}} \cdot a^{\mathrm{d}}_{\mathrm{D}}}{a^{\mathrm{a}}_{\mathrm{A}} \cdot a^{\mathrm{b}}_{\mathrm{B}}} \tag{17.45}$$

or

$$E_{\text{rev}} = E^0 + \frac{RT}{nF} \ln \frac{a_{\text{C}}^{\text{c}} \cdot a_{\text{D}}^{\text{d}}}{a_{\text{A}}^{\text{a}} \cdot a_{\text{B}}^{\text{b}}} \tag{17.46}$$

or, in short and general,

$$E_{\text{rev}} = E^0 - \frac{RT}{nF} \ln \frac{\Pi a_{\text{p}}^{\nu}}{\Pi a_{\text{r}}^{\nu}} \tag{17.47}$$

where Π denotes the product of activities, p and r denote products and reactants, respectively, and ν is a stoichiometric coefficient. This is the Nernst equation. Recalling that $\ln x = 2.3 \log x$, the Nernst equation may also be written as we did in Chapter 5, namely,

$$E_{\text{rev}} = E^0 + \frac{2.3RT}{nF} \log \frac{c_{\text{Ox}}}{c_{\text{Red}}} \tag{5.1}$$

Note that when the product of the activities of the reactants is equal to that of the products, we get from the Nernst equation $E_{\text{rev}} = E^0$. As a good approximation, we can neglect the activity coefficients and substitute concentrations, rather than activities, in the Nernst equation. In addition, at $T = 25\,^{\circ}\text{C}$, $\frac{2.3RT}{F} \cong 0.059$.

The Nernst equation shows that oxidation of a pure metal, $\text{M(s)} \rightarrow \text{M}^{2+}\text{(aq)} + 2\text{e}^-$, will always take place spontaneously in the beginning, from thermodynamics standpoint, because, in this case, $E_{\text{rev}} = E^0 - \frac{RT}{2F} \ln a_{\text{M}^{2+}}$ and $\lim_{x \to 0+} \ln(x) = -\infty$, and thus $\Delta E \rightarrow +\infty$ and $\Delta G \rightarrow -\infty$. Additional discussion of the Nernst equation can be found in Chapter 5.

17.3.5 The Potential–pH (Pourbaix) Diagrams

A very useful method of describing the stability of metals in aqueous solutions is the potential–pH diagrams introduced by Marcel Pourbaix. These equilibrium diagrams relate the reversible potentials of reactions of interest in corrosion studies to the pH and the concentration of different ionic species in solution. We shall use a number of examples to illustrate the principles involved, starting with the most basic diagram relating to water and some of the ionic and molecular species at equilibrium with it. To construct such diagrams, one has to identify the chemical and electrochemical reactions of interest and write the appropriate chemical equilibria and Nernst equations, respectively.

A very simple Pourbaix diagram, for water, appears in Figure 17.4. The shaded area represents the region of thermodynamic stability of water (where both H^+ and OH^- co-exist). Electrolysis cannot occur inside this region. Above and below it, oxygen evolution and hydrogen evolution, respectively, are thermodynamically possible. Whether these reactions will in fact occur at a measurable rate depends on their kinetic parameters.

The HER can be written as

$$2\text{H}^+ + 2\text{e}^- \rightleftharpoons \text{H}_2 \tag{17.4}$$

in acidic solutions, or

$$2\text{H}_2\text{O} + 2\text{e}^- \rightleftharpoons \text{H}_2 + 2\text{OH}^- \tag{17.48}$$

in alkaline solutions.

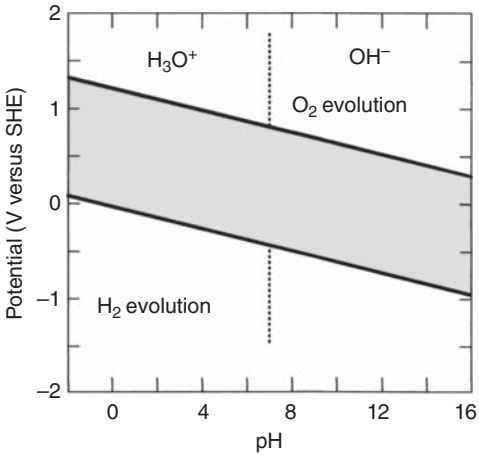


Figure 17.4 The Pourbaix diagram for water. The partial pressures of oxygen and hydrogen are taken as unity. The shaded area is the region of thermodynamic stability of water.

The Nernst equation for Reaction (17.4) is

$$E_{\text{rev}} = 0 + \frac{2.3RT}{2F} \log \frac{[\text{H}^+]^2}{[\text{H}_2]} \tag{17.27}$$

At $T = 25\,^\circ\text{C}$, we get

$$E_{\text{rev}} = \frac{0.059}{2} \log \frac{[\text{H}^+]^2}{[\text{H}_2]} = 0.059 \log[\text{H}^+] - \frac{0.059}{2} \log[\text{H}_2] \tag{17.49}$$

For $[\text{H}_2] = 1$ atm, and substituting Eq. (6.38), we get

$$E_{\text{rev}} = -0.059\,\text{pH} \tag{17.50}$$

the equation of the lower bold line in Figure 17.4. Note that if all stoichiometric coefficients in Eq. (17.4) would be divided by 2, Eq. (17.50) would still be derived. In addition, the HER in alkaline solutions (Eq. (17.48)) yields Eq. (17.50) as well, although the Gibbs free energies for different species are different (namely, $E^0_{\text{H}^+/\text{H}_2} \neq E^0_{\text{H}_2\text{O}/\text{H}_2}$).

The OER can be written as

$$\text{O}_2 + 4\text{H}^+ + 4\text{e}^- \rightleftharpoons 2\text{H}_2\text{O} \tag{17.5}$$

in acidic solutions, or

$$\text{O}_2 + 2\text{H}_2\text{O} + 4\text{e}^- \rightleftharpoons 4(\text{OH})^- \tag{17.6}$$

in alkaline solutions.

The Nernst equation for Reaction (17.5) at 25 °C is

$$E_{\text{rev}} = E^0_{\text{O}_2/\text{H}_2\text{O}} + \frac{0.059}{4} \log([\text{O}_2] \cdot [\text{H}^+]^4) = 1.229 + \frac{0.059}{4} \log[\text{O}_2] - 0.059\,\text{pH} \tag{17.51}$$

The standard potential for the OER was taken from Table 17.1. For $[\text{O}_2] = 1$ atm we get

$$E_{\text{rev}} = 1.229 - 0.059\,\text{pH} \tag{17.52}$$

the equation of the upper bold line in Figure 17.4. The slope in Eq. (17.52) is identical to that in Eq. (17.50). From Eq. (17.51) it is evident that as the partial pressure of oxygen is raised, the reversible potential becomes more positive. Indeed, oxygen is an oxidizing agent. One way to control corrosion is to reduce the partial pressure of oxygen by deaeration. Similarly, hydrogen is a reducing agent; increasing its partial pressure results in shift of the bottom border line down. It can be shown that the neutrality of water from oxidizing/reducing standpoint ($p_{H_2} = 2p_{O_2}$) follows the line

$$E_{rev} = 0.817 - 0.059\,pH \tag{17.53}$$

Oxidizing media prevail above this line, whereas reducing media prevail below it.

The region of stability of water at 25 °C is 1.229 V independent of pH, since the reversible potentials for hydrogen and oxygen evolution change with pH in the same manner. This, incidentally, is the potential region in which the hydrogen/oxygen fuel cell can operate. Thermodynamic considerations lead us to the conclusion that a water electrolyzer must operate at potentials *above* 1.229 V, whereas a hydrogen/oxygen fuel cell must operate at potentials lower than this value.[7] Typical values are 1.6–2.0 V for the former and 0.6–0.8 V for the latter. From the point of view of energy consumption or production, all we need to know is the cell voltage. From the point of view of corrosion, however, we shall see that the potential with respect to the reversible hydrogen electrode in the same solution is very important.

The solid lines bounding the shaded area in Figure 17.4 are the reversible potentials for oxygen and hydrogen evolution as a function of pH. These two lines represent electrochemical equilibria. The dashed vertical line corresponds to the chemical self-ionization shown in Eq. (17.25), at equal concentrations of the two ions. Hence, it is the border between acidic and alkaline media.

The effect of partial pressure of oxygen and hydrogen on the region of stability of water is rather small (see Eqs. (17.28) and (17.29)). For example, increasing the partial pressure of both gases from 1 to 10 atm will increase the potential by 44 mV, from 1.229 to 1.273 V. This, incidentally, can be the basis for the technology of production of hydrogen or oxygen at high pressure by electrolysis of water, without the need to use a gas compressor.

Next, let us focus on several general principles in the construction of Pourbaix diagrams. All electrochemical reactions should be expressed in their reduction form (as we did in Table 17.1), and the equation of the borderline (namely, the reversible potential relative to the SHE) should be calculated accordingly. Reactions that are not dependent on pH (i.e. electrochemical equilibrium in which H^+ and OH^- do not participate) are represented in the Pourbaix diagrams by horizontal lines. In contrast, reactions that are not dependent on potential (i.e. chemical equilibrium in which the oxidation state does not change) are represented in the Pourbaix diagrams by vertical lines. For equilibrium between a solid

7 The thermodynamic limit for a hydrogen–air fuel cell operating at ambient pressure is a little less, because the partial pressure of oxygen in air is only 0.2 atm. However, according to Eq. (17.29) the difference is only 10.3 mV.

and a soluble species, the concentration of the latter must be specified. A series of parallel lines, corresponding to different concentrations, are sometimes drawn. They are noted by the values of −log[species]. It is customary to simplify the diagrams by showing only the lines corresponding to a concentration of 1 μM of each soluble species. This is reasonable, in view of the fact that a moderate rate of corrosion may correspond to about 10 μA cm^{-2} or less, which cannot cause a significant accumulation of soluble corrosion products near the electrode surface, except in confined areas, such as pits or crevices. It is important to carefully list first all species that can be expected in the specific system, as the combination of species taken into account determines the shape of the Pourbaix diagram.

Three possible states of metals are distinguishable in the Pourbaix diagrams. The region of *immunity* is cathodic to the reversible potential of the metal, where it cannot be oxidized. The metal is considered to be totally immune from corrosion attack and safe to use. Cathodic protection (CP) may be used to shift the potential of a metal to this region. The more the immunity region overlaps with the water stability region, the safer the metal is against corrosion attack. Above the reversible potential is the region of *corrosion*. In this region, the metal is stable as an ionic (soluble) product (e.g. Fe^{2+} or Fe^{3+}), and is therefore susceptible to corrosion attack. In the region of *passivity*, the metal can be oxidized anodically, but the product is an insoluble oxide or hydroxide. Whether the metal will actually be passivated in this region depends on the nature of the film formed on its surface, in particular on its porosity and adherence as well as its ionic and electronic conductivity.

Now, let us construct and analyze two Pourbaix diagrams of two metals that are of great interest for construction applications – aluminum and iron. Aluminum is soluble both in acidic and in alkaline media. Its potential–pH diagram is shown in Figure 17.5. The most important feature in this diagram is a passivation region at intermediate pH values, with corrosion possible at both higher and lower pH. The two soluble species are Al^{3+} and AlO_2^-, and the lines representing their equilibria with the various solid phases correspond to a concentration of 1 μM of each metal ion, as explained above.

The first equilibrium to consider is between the two soluble species

$$Al^{3+} + 2H_2O \rightleftharpoons AlO_2^- + 4H^+ \tag{17.54}$$

Since there is no change in valence of the aluminum present in these two ionic species, the associated equilibrium is independent of the potential, and can be expressed for standard conditions by

$$\Delta G^0 = -RT \ln K \tag{17.38}$$

where

$$K = \frac{a_{AlO_2^-} \cdot a_{H^+}^4}{a_{Al^{3+}} \cdot a_{H_2O}^2} \tag{17.55}$$

Assuming that the activity of H_2O is unity (which is correct for a pure solid or liquid; addition of solutes would lower the activity of water) and that the activities of the two ionic species are equal, we obtain

$$K = a_{H^+}^4 \tag{17.56}$$

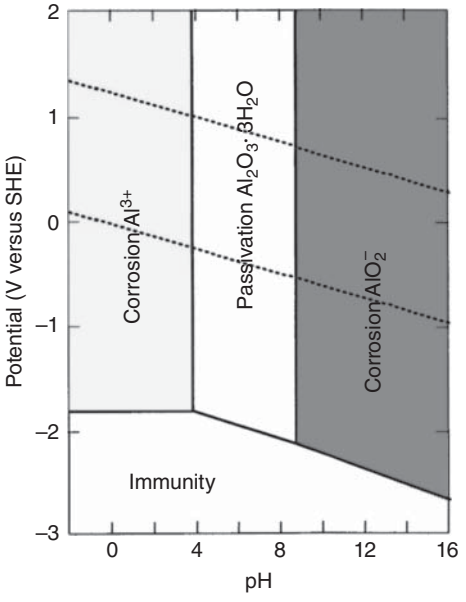


Figure 17.5 Potential–pH diagram for aluminum at 25 °C. The solid phase of the insoluble oxide is assumed to be of the hydrargillite type ($Al_2O_3 \cdot 3H_2O$). Filled areas represent regions where soluble species are stable and, therefore, corrosion can occur from the thermodynamic standpoint. The borderlines correspond to a concentration of 1 μM of each metal ion.

or

$$\log K = -4\,\mathrm{pH} \tag{17.57}$$

But,

$$\log K = -\frac{\sum \nu_\alpha \mu_\alpha^0(\mathrm{cal\,mol^{-1}})}{1363} \quad \text{at } T = 25^\circ\mathrm{C} \tag{17.34}$$

Using the following standard thermodynamic data from the literature, $\mu^0_{Al^{3+}} = -115{,}000\ \mathrm{cal\,mol^{-1}}$, $\mu^0_{AlO_2^-} = -200{,}710\ \mathrm{cal\,mol^{-1}}$, $\mu^0_{H^+} = 0$, $\mu^0_{H_2O} = -56{,}690\ \mathrm{cal\ mol^{-1}}$, we get

$$\log K = -\frac{(-200{,}710 + 4 \times 0) - (-115{,}000 + 2 \times (-56{,}690))}{1363}$$
$$= -\frac{27{,}670}{1363} = -20.30 \tag{17.58}$$

Combining Eqs. (17.57) and (17.58) we get pH = 5.08. This should be represented in the Pourbaix diagram by a vertical line at this pH value, like the equilibrium between H_3O^+ and OH^- in Figure 17.4. In the case of aluminum, this is irrelevant because in the range of about $4.0 \leq \mathrm{pH} \leq 8.6$ only the solid phase is thermodynamically stable.

There are two chemical equilibria in Figure 17.5, represented by the two vertical lines, between the hydrated oxide $Al_2O_3 \cdot 3H_2O$ (hydrargillite) and the two ions, and three electrochemical equilibria between metallic aluminum, the two ions, and the oxide. The equilibrium between metallic Al and Al^{3+} is simple, and its reversible potential is independent of pH, since there are no protons or hydroxyl ions involved:

$$Al^{3+} + 3e^-_M \rightleftharpoons Al^0_M \tag{17.59}$$

The corresponding Nernst equation at 25 °C is

$$E_{\text{rev}} = -\ 1.662 + \frac{0.059}{3}\log[\text{Al}^{3+}] \tag{17.60}$$

For $[\text{Al}^{3+}] = 10^{-6}$ M, we get $E_{\text{rev}} = -1.78$ V. This will be a horizontal line in the Pourbaix diagram.

At intermediate pH values we have the equilibrium

$$\text{Al}_2\text{O}_3 + 6\,\text{H}^+ + 6\,\text{e}_{\text{M}}^- \rightleftharpoons 2\text{Al}_{\text{M}}^0 + 3\text{H}_2\text{O} \tag{17.61}$$

for which the reversible potential at $T = 25$ °C is given by

$$E = -1.550 - 0.059\ \ \text{pH} \tag{17.62}$$

The standard potential was calculated using Eq. (17.38), the standard chemical potentials listed below Eq. (17.34), and $\mu^0_{\text{Al}_2\text{O}_3} = -384{,}530\ \ \text{cal mol}^{-1}$. Thus,

$$E^0 = -\frac{\Delta G^0}{nF} = -\frac{(2\times 0 + 3\times(-56{,}690)) - (-384{,}530 + 6\times 0)}{6\times 23{,}060} = -1.550\,\text{V} \tag{17.63}$$

In alkaline media, the equilibrium to be considered is

$$\text{AlO}_2^- + 2\,\text{H}_2\text{O} + 3\,\text{e}_{\text{M}}^- \rightleftharpoons \text{Al}_{\text{M}}^0 + 4\,\text{OH}^- \tag{17.64}$$

and the appropriate Nernst equation at $T = 25$ °C is

$$E_{\text{rev}} = -1.262 - \frac{0.236}{3}\ \ \text{pH} + \frac{0.059}{3}\ \ \log[\text{AlO}_2^-] \tag{17.65}$$

The standard potential was calculated using Eq. (17.38) in this case too. For $[\text{AlO}_2^-] = 10^{-6}$ M we get

$$E_{\text{rev}} = -1.38 - 0.079\ \ \text{pH} \tag{17.66}$$

Finally, we should consider the two chemical equilibria between $\text{Al}_2\text{O}_3{\cdot}3\text{H}_2\text{O}$ and the aluminum ions. The equilibrium with Al^{3+} is written as

$$2\text{Al}^{3+} + 3\,\text{H}_2\text{O} \rightleftharpoons \text{Al}_2\text{O}_3 + 6\,\text{H}^+ \tag{17.67}$$

Using the values of the standard chemical potentials given earlier we get

$$\log\ K = -\frac{(-384{,}530 + 6\ \times\ 0) - (2\ \times\ (-115{,}000) + 3\ \times\ (-56{,}690))}{1363} = -\frac{15{,}540}{1363} = -11.40 \tag{17.68}$$

In addition,

$$K = \frac{[\text{H}^+]^6}{[\text{Al}^{3+}]^2} \tag{17.69}$$

Therefore,

$$-11.40 = -6\,\text{pH} - 2\,\log[\text{Al}^{3+}] \tag{17.70}$$

For $[\text{Al}^{3+}] = 10^{-6}$ M, we get pH = 3.9.

Finally, the chemical equilibrium between $Al_2O_3 \cdot 3H_2O$ and AlO_2^- is written as

$$Al_2O_3 + 2OH^- \rightleftharpoons 2AlO_2^- + H_2O \tag{17.71}$$

Using the values of the standard chemical potentials given earlier and $\mu^0_{OH^-} = -37{,}570 \ \text{cal mol}^{-1}$, we get

$$\log K = -\frac{(2 \times (-200{,}710) + (-56{,}690)) - (-384{,}530 + 2 \times (-37{,}570))}{1363}$$
$$= -\frac{1560}{1363} = -1.145 \tag{17.72}$$

In addition,

$$K = \frac{[AlO_2^-]^2}{[OH^-]^2} \tag{17.73}$$

Therefore,

$$-1.145 = 2 \log[AlO_2^-] - 2 \log[OH^-] \tag{17.74}$$

For $[AlO_2^-] = 10^{-6}$ M, we get

$$-\log[OH^-] = 5.4275 \tag{17.75}$$

which can be rewritten as pH = 8.57.

Aluminum represents an interesting case, which warrants further discussion. We note that the limit of the region of thermodynamic immunity lies at very negative potentials with respect to the lower limit of stability of water (which is the reversible potential for hydrogen evolution) at all pH values. Thus, one would expect rapid dissolution of this metal in any aqueous medium.[8] This is indeed found in acidic and alkaline solutions, but around neutral pH the oxide formed is very dense and nonconducting, and oxidation is effectively stopped after a thin layer of about 5 nm has been formed. This thin layer of oxide permits aluminum to be used as a construction material and in many other day-to-day applications. There are, of course, additional ways (e.g. anodizing and painting) to protect aluminum that is exposed to harsh environments, beyond the protection afforded by the spontaneously formed oxide film. However, the unique feature of this metal (and several others; e.g. titanium, tantalum, and niobium) is that it repassivates spontaneously when the protective layer is removed mechanically or otherwise, as long as the pH of the medium in contact with it is in the appropriate range shown in Figure 17.5. The aluminum susceptibility to corrosion in both acidic and basic environments is referred to as an amphoteric behavior. From Figure 17.5 it is also evident that very high cathodic potentials are required in order to reduce aluminum ions and oxides to their metallic form. Therefore, electroplating of aluminum is *not* carried out in industry from aqueous electrolyte solutions.

Next, we consider the Pourbaix diagram for iron, which is, of course, of paramount importance for the understanding of corrosion of ferrous alloys such

8 This statement is a little careless, since we cannot deduce the rate of a reaction from thermodynamic data alone. Yet, when there is a very large driving force (i.e. when the system is far from equilibrium), the reaction will tend to be fast, unless some special mechanism prevents it or slows it down.

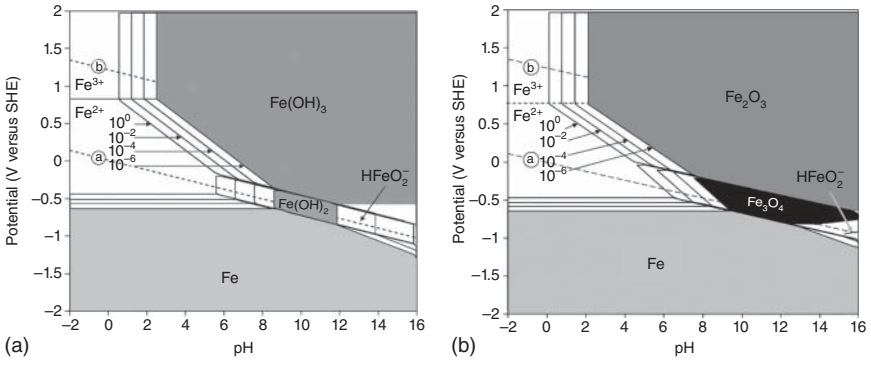


Figure 17.6 The potential–pH diagram for iron at 25 °C. The lines for four concentrations of soluble species are drawn. (a) Wet (hydrous) corrosion products. (b) Dry (anhydrous) corrosion products. Source: Reproduced with permission from P.R. Roberge, *Corrosion Engineering: Principles and Practice*, McGraw-Hill, 2008.

as the many types of steel. Figure 17.6a illustrates the E–pH diagram for iron at 25 °C *in the presence of water or in humid environments* (i.e. the hydrous form). This diagram was constructed by considering all possible reactions associated with iron in wet or aqueous conditions, excluding dry forms of corrosion products such as magnetite (Fe_3O_4) or iron (ferric) oxide (Fe_2O_3), both of which are important iron ore constituents. The various stability regions for these drier (anhydrous) corrosion products are shown in Figure 17.6b, where the predominant compounds and ions are also indicated. At potentials more positive than −0.6 V versus SHE and at pH values below 9, ferrous ion (Fe^{2+}) is the stable substance. This indicates that iron will corrode under these conditions. In other regions of the iron Pourbaix diagram, it can be seen that the corrosion of iron produces ferric ions (Fe^{3+}), ferric hydroxide ($Fe(OH)_3$), and ferrous hydroxide ($Fe(OH)_2$). At pH ≥ 12.5, the anion $HFeO_2^-$ is thermodynamically stable, showing that iron can also be amphoteric to some extent.

The region below pH 9 includes most natural environments with which structural materials are commonly in contact, making iron and many of its alloys vulnerable to corrosion. There are no soluble species between pH 9 and 12.5, and iron could pass directly from the immune to the passive region as the potential is increased (e.g. due to anodic protection). Also, it should be possible to passivate iron in solutions where the pH exceeds about 4, by oxidizing the surface, either chemically or electrochemically. In this case, the potential can be forced to cross the active region (where Fe^{2+} is the stable species) rapidly, to reach the region of passivity, where $Fe(OH)_3$ is stable.

The presence of a relatively large immunity region in Figure 17.6 indicates that iron may corrode much less under these potential–pH conditions. These diagrams also indicate that if the potential of iron is made sufficiently negative or shifted cathodically below approximately −0.5 V versus SHE in neutral or acidic environments (e.g. by connecting the iron to zinc or magnesium, which serve as sacrificial anodes), iron will corrode much less. This explains the generally

accepted cathodic protection criterion of −0.53 V versus SHE used across industries to protect steel assets buried in soils. One should note that at this potential, water electrolysis into hydrogen is likely to occur.

Let us show how the Pourbaix diagrams of iron were constructed. The electrochemical equilibrium between Fe(s) and Fe^{2+} can be expressed by

$$Fe^{2+} + 2e^- \rightleftharpoons Fe \tag{17.76}$$

Using the Nernst equation and Table 17.1 we can write

$$E_{rev} = -0.440 + \frac{0.059}{2} \log[Fe^{2+}] \tag{17.77}$$

For $[Fe^{2+}] = 10^{-6}$ M, we get $E_{rev} = -0.62$ V. This is a horizontal line.

The electrochemical equilibrium between Fe^{3+} and Fe^{2+} is expressed by

$$Fe^{3+} + e^- \rightleftharpoons Fe^{2+} \tag{17.78}$$

Again, using the Nernst equation and Table 17.1 we can write

$$E_{rev} = +0.771 - 0.059 \log \frac{[Fe^{2+}]}{[Fe^{3+}]} \tag{17.79}$$

For equal ion activities we get $E_{rev} = +0.771$ V. This is also a horizontal line.

The chemical equilibrium between Fe^{2+} and $Fe(OH)_2$ (or its equivalent $FeO{\cdot}H_2O$, white oxide) is a precipitation reaction that can be written for acidic media as

$$Fe^{2+} + 2H_2O \rightleftharpoons Fe(OH)_2 + 2H^+ \tag{17.80}$$

or for alkaline media as

$$Fe^{2+} + 2OH^- \rightleftharpoons Fe(OH)_2 \tag{17.81}$$

Using the definition of the precipitation constant as in Eq. (17.32) and a value of $K_{sp} = 10^{-14.71}$ we can write

$$10^{-14.71} = [Fe^{2+}] \cdot [OH^-]^2 \tag{17.82}$$

But, at 25 °C

$$\log[OH^-] = pH - 14 \tag{17.83}$$

Thus,

$$-14.71 = \log[Fe^{2+}] + 2pH - 28 \tag{17.84}$$

or

$$pH = 6.645 - \frac{1}{2} \log[Fe^{2+}] \tag{17.85}$$

For $[Fe^{2+}] = 10^{-6}$ M, we get pH = 9.645.

The electrochemical equilibrium between Fe^{2+} and Fe_3O_4 can be represented by

$$Fe_3O_4 + 8H^+ + 2e^- \rightleftharpoons 3Fe^{2+} + 4H_2O \tag{17.86}$$

Using the Nernst equation we can write

$$E_{rev} = E^0_{Fe_3O_4|Fe^{2+}} - \frac{0.059}{2} \log \frac{[Fe^{2+}]^3}{[H^+]^8} \tag{17.87}$$

The standard potential of this reaction can be calculated using Eq. (17.35), substituting the values $\mu^0_{Fe_3O_4} = -242{,}400$ cal mol^{-1}, $\mu^0_{Fe^{2+}} = -20{,}300$ cal mol^{-1}, $\mu^0_{H^+} = 0$, $\mu^0_{H_2O} = -56{,}690$ cal mol^{-1}. Thus, we get

$$E^0_{Fe_3O_4|Fe^{2+}} = -\frac{(3 \times (-20{,}300) + 4 \times (-56{,}690)) - (-242{,}400 + 8 \times 0)}{23{,}060 \times 2}$$
$$= +0.981\,\text{V} \tag{17.88}$$

Substituting Eq. (17.88) into Eq. (17.87) and rearranging we get at 25 °C:

$$E_{rev} = 0.981 - 0.0885 \log[Fe^{2+}] - 0.236\ \text{pH} \tag{17.89}$$

The electrochemical equilibrium between Fe^{2+} and Fe_2O_3 can be represented by

$$Fe_2O_3 + 6H^+ + 2e^- \rightleftharpoons 2Fe^{2+} + 3H_2O \tag{17.90}$$

Because the ratio between the stoichiometry coefficients of H^+ to e^-, 3 : 1, is smaller than in Eq. (17.86), we can expect the slope of the line that represents Eq. (17.90) to be smaller than that of Eq. (17.86). Using the Nernst equation we can write

$$E_{rev} = E^0_{Fe_2O_3|Fe^{2+}} - \frac{0.059}{2} \log \frac{[Fe^{2+}]^2}{[H^+]^6} \tag{17.91}$$

We can use the standard chemical potentials $\mu^0_{Fe_2O_3} = -177{,}100$ cal mol^{-1} and $\mu^0_{Fe(OH)_3} = -161{,}930$ cal mol^{-1}, along with the ones listed before, to calculate the standard potentials of the reduction reactions forming either the hydrous hydroxide or the anhydrous red oxide (note that $2Fe(OH)_3$ is equivalent to $Fe_2O_3 + 3H_2O$). Thus, we get

$$E^0_{Fe_2O_3|Fe^{2+}} = -\frac{(2 \times (-20{,}300) + 3 \times (-56{,}690)) - (-177{,}100 + 6 \times 0)}{23{,}060 \times 2}$$
$$= +0.728\,\text{V} \tag{17.92}$$

$$E^0_{Fe(OH)_3|Fe^{2+}} = -\frac{(2 \times (-20{,}300) + 3 \times (-56{,}690)) - (-161{,}930 + 6 \times 0)}{23{,}060 \times 2}$$
$$= +1.057\,\text{V} \tag{17.93}$$

Hence, the Nernst equations for the hydrous and anhydrous forms are

$$E_{rev,anhydrous} = 0.728 - 0.1773\,\text{pH} - 0.059 \log[Fe^{2+}] \tag{17.94}$$

$$E_{rev,hydrous} = 1.057 - 0.1773\,\text{pH} - 0.059 \log[Fe^{2+}] \tag{17.95}$$

We shall conclude this section by making some general remarks on the advantages and limitations of potential–pH diagrams. It has already been stated that these are equilibrium diagrams. Hence, we can learn from them what *cannot*

happen (e.g. a metal cannot be anodically dissolved in the region in which it is "immune," and water cannot be electrolyzed in the region of its stability, which is indicated in all such diagrams). We cannot deduce which reaction *will happen* at a measurable rate, because these diagrams do not reflect the kinetics. The fact that at a certain pH and potential a metal can corrode according to its Pourbaix diagram is no proof that it actually will do so. The regions marked as "passivation" only indicate that a solid iron compound (usually an oxide or a hydroxide) is the thermodynamically stable corrosion product. Whether passivation will or will not take place depends on the nature of the oxide and on the environment in contact with it.

An important point to remember is that potential–pH diagrams are usually given for pure elements. Now, high-purity metals belong to the research laboratory, but nothing is ever constructed of a pure metal. In fact, the most important developments in the field of metallurgy have been in the area of new and improved alloys, designed to fit specific engineering requirements. The corrosion behavior of an alloy is rarely, if ever, a linear combination of the corrosion behavior of its components. Even for a given composition, the corrosion of an alloy usually depends on metallurgical factors such as the grain size and heat treatment of the material. An extreme example is the high corrosion resistance of some of the so-called *glassy metals* or *amorphous alloys*, compared to crystalline alloys of the same composition.

Pourbaix diagrams also have other limitations: (i) they typically refer to pure water, but in practice there are other ions that might affect equilibria. (ii) As explained before, their shape is affected by the species that are taken into account; thus, previous knowledge of the specific system is required. (iii) The pH is that in the vicinity of the metal surface, which could differ significantly from the pH of the bulk solution.

In spite of the foregoing limitations, the corrosion scientist and engineer can derive a wealth of information by consulting the relevant potential–pH diagrams. The regions of immunity, passivity, and corrosion are demarcated, and the most common corrosion products are shown. The Pourbaix diagrams can be used to study the effect of inhibitors on the corrosion of a metal in an aqueous solution, roughly determine the right conditions for electroplating, study the equilibrium of minerals, etc. Studying the relevant diagram is an excellent way to start a new corrosion study, but it should never be the only tool used to solve the problem.

17.4 Kinetics of Corrosion

17.4.1 Introduction and Important Terms

Understanding the kinetics of corrosion requires proper definition and discussion of terms such as exchange current density, the Tafel equation, polarization, the Butler–Volmer equation, and the Mixed-Potential Theory. Some of these have already been discussed in previous chapters; for example, electrode kinetics, the relation of potential to the energy of activation, and the symmetry factor (Chapter 4), overpotential, the Transition-State Theory, the Butler–Volmer equation and the empirical Tafel equation (Chapter 5), the transfer coefficient

(Chapter 6), and the EDL (Chapter 8). Therefore, these will not be discussed in this section at all, or will be discussed only briefly.

Polarization is the change in potential, relative to the equilibrium potential, that is caused by the passage of current. The deviation from the equilibrium value results mainly from resistance polarization, activation polarization, and concentration polarization. Knowing the kind of polarization that occurs can be very helpful, since it allows an assessment of the determining characteristics of a corroding system. For example, if corrosion is controlled by concentration polarization, then any change that increases the diffusion rate of the active species (e.g. oxygen) will also increase the corrosion rate. In such a system, it would therefore be expected that agitating the liquid or stirring it would tend to increase the corrosion rate of the metal. However, if a corrosion reaction is activation controlled, then stirring or increasing the agitation will have no effect on the corrosion rate. With respect to *activation polarization*, the exchange current density, j_0, depends on the temperature, redox reaction, concentration of both reactants and products, as well as on the composition, surface contamination, and surface roughness of the solid electrode. When j_0 is very high (say, at least 1 mA cm^{-2}), the reaction is considered reversible. When it is low, the kinetics is slow, impurities might affect the results, and the activation overpotential η_{ac} has to be higher in order to get the same net current. We have shown that the value of the exchange current density for the HER on different metals can be orders of magnitude different, from 10^{-11} to 10^{-2} A cm^{-2} (note, however, that for each metal there is a wide range of values reported in the literature, depending, among others, on the nature of the electrolyte, pH, the crystal face, and the preparation of the electrode surface). As we will show later in this section, this can be the major reason for the difference in the corrosion rates of different metals in acidic media. With respect to *concentration polarization*, the solubility of oxygen in air-saturated water is very low, namely 8.29 ppm (or 259.0 μM) at 25 °C. Therefore, a frequent case of concentration polarization occurs when the cathodic processes depend on the reduction of dissolved oxygen. With respect to *resistance polarization*, the resistivity of water depends on its quality, from 18 MΩ cm for pure water, down to 30 Ω cm for typical seawater (coastal). Therefore, the corrosion rate may change significantly depending on the type of water the metal is exposed to.

Next, let us refer to the application of an anodic overpotential and illustrate schematically the essence of the Transition-State Theory (also known as the Activated-Complex Theory) and the symmetry factor. This theory was first developed by Henry Eyring in 1935, and was later developed by Polanyi, Evans, and others, to explain the reaction rates of chemical reactions. Until then, the empirical Arrhenius rate law was widely used to determine reaction barrier energies. The Transition-State Theory assumes quasi-equilibrium between reactants and activated transition-state complexes: $A + B \rightleftharpoons [AB^{\ddagger}] \rightarrow$ products. The ‡ notation refers to the value of interest at the transition state. The reaction rate is then based on statistical mechanics of this equilibrium, and is directly proportional to the concentration of this complex multiplied by the frequency at which it is converted into a product. The reaction coordinates diagram describes energy changes during the course of the reaction; the reactants and products lie at energy minima and are separated by an energy barrier, the maximum of

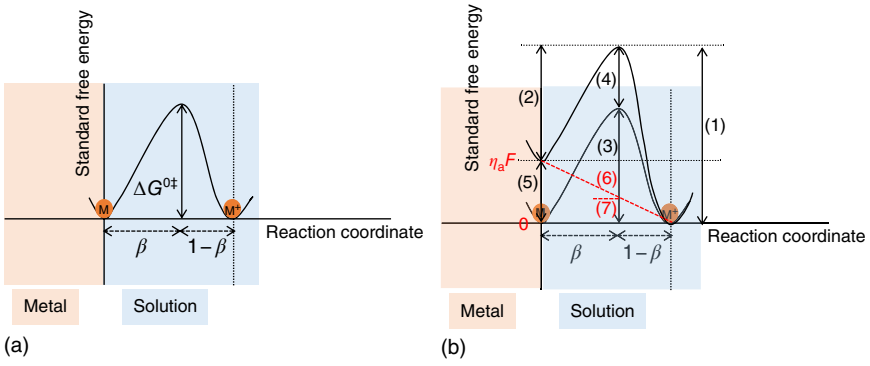


Figure 17.7 Schematic representation of the standard free energy change during a faradaic reaction. (a) At equilibrium. (b) At a more positive applied potential (anodic polarization). β is the symmetry factor and η_a is the anodic overpotential. (1) $\Delta\overline{G}_c^{0\ddagger}$, (2) $\Delta\overline{G}_a^{0\ddagger}$, (3) $\Delta\overline{G}^{0\ddagger}$, (4) $(1-\beta)F\eta_a$, (5) $F\eta_a$, (6) $\beta F\eta_a$, and (7) $(1-\beta)F\eta_a$. The structure of the EDL is shown in Figure 8.5.

which is called the transition state. The energy curves around the minima are drawn as parabolic shapes, following the harmonic oscillation approximation. Now, let us refer to the reaction $M(s) \rightleftharpoons M^+(aq) + e^-$. The metal atoms on the electrode surface are in energy wells associated with the lattice structure, and in order to pass into the solution they have to overcome an energy barrier (activation energy). The structure of the EDL, arrangement of ions in the vicinity of the electrode, and the electrode potential drop across the EDL were already described in Chapter 8 and Figure 8.5 therein. Figure 17.7a illustrates the equilibrium state in which the barrier of the forward (anodic) reaction is the same as the barrier of the backward (cathodic) reaction. At equilibrium, the flux of charge through the double layer is the same in both directions, and the net current is zero. If we apply positive (anodic) overpotential, the energy well of atoms at the metal surface increases by $F\eta_a$ (for the transfer of one electron; see Figure 17.7b). As a result of this anodic polarization, a net anodic current will be established.

β is a measure of the symmetry of the energy barrier. When $\beta = 0.5$, the degree of decrease in the free energy barrier in the forward direction is the same as the degree of increase in the reverse direction. When $\beta < 0.5$, the slope of the parabolic energy curve for the product is steeper than for the reactant. When $\beta > 0.5$, the slope of the parabolic energy curve for the reactant is steeper than for the product. If $\beta = 0$, none of the applied energy will promote the anodic reaction, i.e. $\Delta\overline{G}_a^{0\ddagger} = \Delta G^{0\ddagger}$ (energy barriers (2) and (3) are the same, and the two maxima will be separated accordingly). At the other extreme, $\beta = 1$, all of the additional energy goes into promoting the anodic reaction, the two maxima coincide, $\Delta\overline{G}_a^{0\ddagger} < \Delta G^{0\ddagger}$, but $\Delta\overline{G}_c^{0\ddagger}$ will be the same.

The Transition-State Theory can be used to derive the Butler–Volmer equation, as we did in Chapter 5:

$$\frac{j}{j_0} = \exp\left(\frac{\beta F}{RT}\eta\right) - \exp\left[-\frac{(1-\beta)F}{RT}\eta\right] \tag{5.25}$$

It is recalled that in deriving Eq. (5.25) it was tacitly assumed that the concentration at the surface is independent of the current density; in other words, mass-transport limitation was neglected. For both the anodic (oxidation) reaction and the cathodic (reduction) reaction, the current density increases exponentially with increasing absolute value of the overpotential, $|\eta|$. The current density in Eq. (5.25) is the net (applied or measured) current density in an external circuit. Equation (5.25) is the *Butler–Volmer equation*, named after the chemists John Alfred Valentine Butler and Max Volmer. Equation (5.25) is valid only for activation-controlled reactions. A plot of the current–potential relation is commonly referred to as a *polarization curve*. Although in most cases the detailed mechanism is not known, Eq. (5.25) can still be used as a fitting function that is valid over a restricted range. Note that Eq. (5.25) was derived in Chapter 5 for an oxidation reaction, using β in the anodic (forward) direction and $(1-\beta)$ in the cathodic (backward) reaction, and η is substituted with its sign (+ for anodic overpotential, − for cathodic overpotential) so that a positive overpotential results in a positive net current and a negative overpotential results in a negative net current. The use of the transfer coefficient instead of the symmetry factor in multistep electrode reactions has already been explained in Chapter 6; therefore, it will not be discussed here.

17.4.2 Two Limiting Cases of the Butler–Volmer Equation: Tafel Extrapolation and Polarization Resistance

In Section 5.2.4 we discussed the two limiting cases (approximations) of Eq. (17.95). Both are very useful in corrosion study. On one hand, if the overpotential is small (up to ±20 mV, depending on the value of the symmetry factor), we can apply the low-field approximation $e^x = 1 + x + x^2/2 \ldots = 1 + x$, for $x \ll 1$, thus obtaining a linear relation between η and i_{net} around E_{rev}:

$$\frac{j}{j_0} = \frac{F}{RT}\eta \tag{5.6}$$

This equation is shown for $n = 1$. The charge-transfer resistance is then defined as

$$R_{\text{ct}} \equiv \frac{\eta}{j} = \frac{RT}{j_0 F} \tag{17.96}$$

Note that R_{ct} is the slope of the E versus j curve, and not E versus $\log j$. The charge-transfer resistance is not constant, but rather depends on the applied current, j_{app}. It is related to a single, kinetically controlled electrochemical reaction at equilibrium, without any mixed potential and at constant surface concentrations. In the literature, R_{ct} is sometimes referred to as identical to the polarization resistance, R_{p}, but it is wrong. The polarization resistance is the slope of the E versus j curve at steady state, around a mixed potential. In this case, the open-circuit potential (OCP) equilibrates at the potential where the cathodic and anodic currents are equal. Therefore, instead of dealing with $\eta = E_{\text{app}} - E_{\text{rev}}$, we deal with $\Delta E = E_{\text{app}} - E_{\text{OCP}}$. The polarization resistance is measured between the working electrode and a reference electrode in solution; thus, it may encompass various

resistive elements beyond R_{ct}, and it is therefore usually larger than R_{ct}. Polarization resistance measurements are used to determine the corrosion current density, using the following equation:

$$j_{corr} = \frac{1}{R_p} \frac{b_a b_c}{2.3(b_a + b_c)} \tag{17.97}$$

where b_a and b_c are the Tafel slopes of the anodic and cathodic reactions, respectively, in units of V decade^{-1}, and R_p is in units of $\Omega\,cm^2$. The Stern–Geary coefficient (SGC), in units of V, is then defined as

$$SGC = \frac{b_a |b_c|}{2.3(b_a + |b_c|)} \tag{17.98}$$

ASTM G59 standard describes the measurements of R_p and the rate of general (uniform) corrosion.

Let us now move to the other limiting case of Eq. (5.25), which results in the Tafel extrapolation. The definition and calculation of the Tafel slopes have already been discussed in Sections 5.2.1, 5.2.4, 6.1.1, and 6.1.3. For a high overpotential, one of the terms in Eq. (5.25) becomes negligible. This will happen when the backward reaction will contribute less than 1% of the current, for example,

$$\frac{\exp\left[-(1-\beta)\frac{F\eta}{RT}\right]}{\exp\left(\frac{\beta F\eta}{RT}\right)} = \exp\left(-\frac{F\eta}{RT}\right) \leq 0.01 \tag{17.99}$$

In other words, this will happen mathematically when $|\eta| > 118$ mV at 25 °C. The Tafel equation is then derived as

$$\eta = a \pm b \log|j| \tag{17.100}$$

where $a = -\frac{2.3RT}{\alpha F} \log j_0$, $b = \frac{2.3RT}{\alpha F}$, the (+) sign is used for anodic reactions, and the (−) sign is used for cathodic reactions.

The Tafel equation is applicable only when the overpotential results from kinetic limitations. It does not exist in the presence of significant voltage drops in the solution. If electrode kinetics is fast, the system will reach j_L before sufficiently high η is attained; therefore, no Tafel relations will be observed (because they require no mass-transport limitations on the current). In contrast, when electrode kinetics is slow and significant η_{ac} is required, Tafel regions may be easily identified. Therefore, a Tafel behavior is indicative of a completely irreversible kinetics. Systems with this property do not allow flow of significant current, except at high η, where the faradaic process is effective in one direction, and is therefore irreversible from the chemical standpoint.

The values of b are usually in the range of 30–300 mV decade^{-1} (the meaning of decade is, for example, that a step from 1 to 100 equals two decades). When b_a and b_c are not known, $b_a = |b_c| = 118$ mV decade^{-1} is used as a default value. For this value and $\eta = \pm 118$ mV, the rate of the anodic and cathodic reactions increase by 1 order of magnitude. A few examples of typical experimental values of the Tafel slopes: $|b_c| = 30$ and 150 mV decade^{-1} for the HER in 1 M HCl on Pt and on Fe, respectively; $b_a = 39$–60 mV decade^{-1} for oxidation of iron in 0.26 M H_2SO_4.

17.4.3 Corrosion Rate

The amount of metal removed by corrosion can be determined from Faraday's law:

$$w = \frac{j_{\text{corr}} AtM}{nF} \tag{17.101}$$

where w is the corroded mass (g), j_{corr} is the corrosion current density (A cm^{-2}), M is the atomic mass of the metal (g mol^{-1}), n is the number of electrons transferred per atom/molecule, t is time (s), F is Faraday's constant (C mol^{-1}), and A is the exposed surface area (cm^2). Possible causes of deviation from Faraday's law include the following: (i) parasitic processes consume some charge; (ii) not all of the reactants are consumed; (iii) the postulated process is not the actual process; and (iv) some of the material from the sample falls of.

The corrosion intensity (CI) can be similarly expressed in units of g cm^{-2} s^{-1} as

$$\text{CI} = \frac{j_{\text{corr}} M}{nF} \tag{17.102}$$

We can also express the corrosion penetration rate (CPR) in units of loss-in-dimension perpendicular to the corroding surface per unit time (mpy):

$$\text{CPR} = \frac{j_{\text{corr}}}{25.4F\rho} \times \frac{M}{n} \times 3.154 \times 10^{11} \tag{17.103}$$

In this equation, the density ρ is substituted in units of g cm^{-3}. The time coefficient takes into account 365 days per year. If we multiple Eq. (17.103) by 25.4 we get the CPR in units of micrometers per year. In Eq. (17.103) we separated M/n because it is the equivalent weight (EW). The number of equivalents, N_{EQ}, is simply

$$N_{EQ} = \frac{1}{\text{EW}} = \sum \left(\frac{f_i}{M_i/n_i} \right) = \sum \left(\frac{f_i n_i}{M_i} \right) \tag{17.104}$$

where f_i is the mass fraction of element i in the alloy. This allows us to calculate the corrosion rate of common alloys (see ASTM G102-89(2015)).

What is an acceptable corrosion rate? This would depend on the application, the possible effect of corrosion on the functionality of the part, safety concerns, and the cost of the material. For example, for iron- and nickel-based alloys, CPR < 1 mpy is exceptional, CPR = 20–50 mpy is reasonable, and CPR > 200 mpy is often unacceptable. The latter may be allowed for cast iron or a very high wall thickness. In contrast, 5–20 mpy may be unacceptable for precious alloys. For medical implants we usually permit less than 1 μm yr^{-1} (0.0394 mpy).

17.4.4 The Mixed-Potential Theory and the Evans Diagrams

The Mixed-Potential Theory was presented by Wagner and Traud in 1938. However, it was not applied until 1950, when Stern applied it in an analysis of corrosion. This theory is based on the following hypotheses: (i) all corrosion reactions involve at least one partial oxidation reaction and one partial reduction reaction. (ii) During the corrosion of an electrically isolated metal, the total partial

anodic current must be equal to the total partial current, because there can be no net accumulation of electric charge on the surface of the electrode. It should be emphasized that here we refer to equation of currents, and not current densities. (iii) For metals, the electric potential of the metal at the anodic site is equal to that at the cathodic site due to their very low resistivities. This means that their OCPs polarize to the common free corrosion potential, E_{corr}, and corrosion current density, j_{corr}. This hypothesis may not be valid for semiconductors, particularly if the anodic and cathodic sites are widely separated.

Let us demonstrate the above for the case of iron corrosion in an acid. When a piece of iron is immersed in 1.0 M HCl it dissolves readily. At what potential will this process occur? The reversible potentials for the anodic and cathodic reactions depend on the composition of the solution in contact with the electrode surface. For the purpose of the present calculation we shall assume that after a short time, the solution is saturated with molecular hydrogen and the concentration of iron is 1 μM. With these assumptions, we have for the anodic reaction

$$E_{rev} = E^0 + \frac{2.3RT}{nF} \log \frac{[Fe^{2+}]}{[Fe^0]} = -0.617 \quad \text{V versus SHE} \tag{17.105}$$

For the cathodic reaction, by definition $E^0 = 0.000$; hence,

$$E_{rev} = E^0 - \frac{2.3RT}{F}\text{pH} = -\frac{2.3RT}{F}\text{pH} \quad \text{V versus SHE} \tag{17.106}$$

Clearly, for anodic dissolution to occur, the potential must be positive with respect to the reversible potential for the Fe^{2+}/Fe couple and negative with respect to the reversible potential for the HER. Thus, all we could predict from this thermodynamic argument is that at pH = 0 the corrosion potential must be somewhere between −0.617 and 0.000 V versus SHE. The rest depends on kinetics. To proceed, we need to know the exchange current densities and the Tafel slopes for the two reactions concerned. The two partial current densities are plotted in Figure 17.8 as a function of potential. The potential must settle at the point where the anodic and cathodic currents are equal, which is called the *corrosion* (or *mixed*) *potential*, E_{corr}. As evident from Figure 17.8, the extrapolations of the anodic and cathodic curves are constructed with the Tafel equation on an E versus $\log j$ diagram. This scheme is called the *Evans diagram*, after the name of the British electrochemist Ulick R. Evans.

For the example shown in Figure 17.8, $E_{corr} = -0.52$ V versus SHE and $j_{corr} = 27\ \text{mA cm}^{-2}$. This is quite a high current density, corresponding to fast dissolution of the metal and vigorous hydrogen evolution, which are indeed observed when iron is dipped into a 1.0 M solution of HCl.

It should be noted that the values of the two Tafel slopes have to be measured, and they may deviate from the values given in the caption of Figure 17.8. If that is the case, the corrosion rate and mixed potential should be decided by the actual values measured. In particular, once an additive is added, the Tafel slope for either anodic or cathodic reactions, or both, may be different.

We can calculate the corrosion potential and the corrosion current in a straightforward manner by writing the Tafel equations for the two partial reactions and

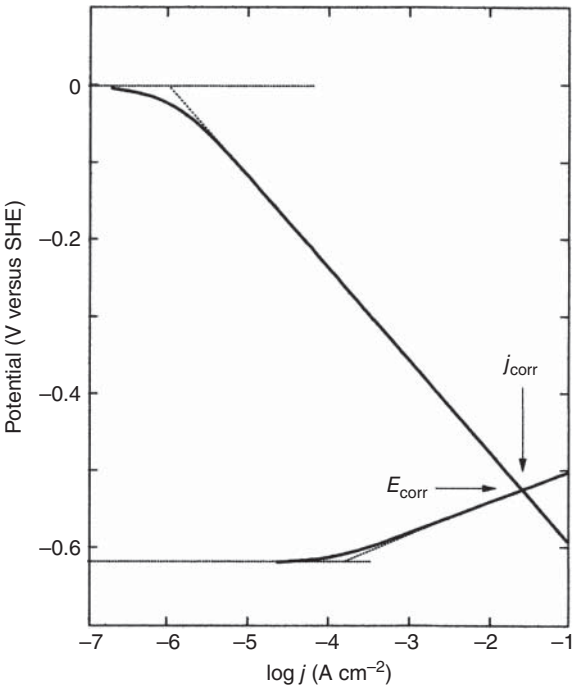


Figure 17.8 Evans diagram showing the currents for iron dissolution and hydrogen evolution at pH = 0, and the resulting values of E_{corr} and of j_{corr}. Parameters for the anodic and cathodic reactions are $j_{0,a} = 1 \times 10^{-4}$ A cm^{-2}, $j_{0,c} = 1 \times 10^{-6}$ A cm^{-2}, $b_a = 0.039$ V decade^{-1}, $b_c = 0.118$ V decade^{-1}, $E_{rev,a} = -0.617$ V versus SHE, $E_{rev,c} = 0.000$ V versus SHE.

solving for the potential at which the currents are equal:

$$j_{0,a} \exp\left[\frac{E_{corr} - E_{rev,a}}{b_a}\right] = j_{0,c} \exp\left[-\frac{E_{corr} - E_{rev,c}}{b_c}\right] \tag{17.107}$$

Now, consider the effect of pH. Assuming that all the kinetic parameters are unchanged[9] and only the reversible potential for the HER is affected, we obtain the curves shown in Figure 17.9. As the pH is increased, the corrosion potential becomes less positive, approaching the reversible potential for the Fe^{2+}/Fe redox couple, and the corrosion current decreases, from $j_{corr} = 27$ mA cm^{-2} at pH = 0 to $j_{corr} = 0.14$ mA cm^{-2} at pH = 6. This is hardly surprising. It simply confirms the common observation that iron corrodes faster in concentrated than dilute acid. The interesting new insight we can gain from Figure 17.9 is that this difference is not directly related to the rate of metal dissolution in the different media. It is, in fact, determined by the different rates of hydrogen evolution, which is necessary to consume the electrons released in the process of oxidizing the metal to its ions. The process is evidently *cathode limited*, and the corrosion potential is close to the reversible potential for the *anodic process*. Since the process is cathode limited, it is possible to slow it down by inhibiting the rate of hydrogen evolution. Many commercial corrosion inhibitors function in this manner. Considering Figure 17.9, it is easy to see that decreasing the exchange current

9 This assumption is made here only for convenience of presentation. Usually, the exchange current density for the HER is found to be lower at intermediate pH values than in strongly acidic or strong alkaline solutions.

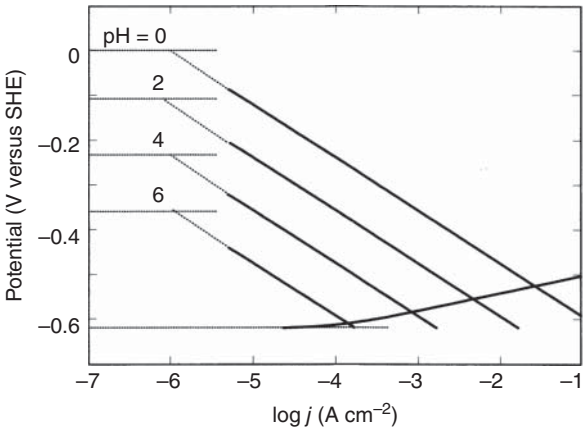


Figure 17.9 Evans diagram showing the current densities for iron dissolution and hydrogen evolution at different pH values. The kinetic parameters, which are the same as in Figure 17.8, have been arbitrarily assumed to be independent of pH.

density of hydrogen evolution by the addition of a suitable corrosion inhibitor is equivalent to increasing the pH, in terms of its effect on E_{corr} and on j_{corr}.

The addition of an oxidizing agent to an acid solution can affect the corrosion potential and current density in various ways, depending on its redox potential and its specific exchange current density. Figure 17.10 illustrates the effect of ferric (Fe^{3+}) salt addition to the solution. Note that according to Table 17.1, the standard potential of the half-cell reaction $Fe^{3+} + e^- \rightleftharpoons Fe^{2+}$ is +0.771 V. The new corrosion potential is established at the point where the total reduction currents ($I_{Fe^{3+}\rightarrow Fe^{2+}} + I_{H^+\rightarrow H_2}$) equal the total oxidation current ($I_{M\rightarrow M^+}$). Thus, the addition of the oxidizer to the solution shifted the corrosion potential in the noble (positive) direction to E^*_{corr}, increased the corrosion rate from j_{corr} to j^*_{corr}, and decreased the hydrogen evolution from j_{corr} to $j_{H^+\rightarrow H_2}$. Note that if the exchange current density of the ferric–ferrous half-cell on metal M was much lower, i.e. its cathodic Tafel slope would have shifted in parallel to the left, the total reduction current would have hardly been affected; therefore, the corrosion potential and corrosion current density would not change either.

In Section 17.4.1 we saw that at 25 °C the exchange current density for the HER on Fe is $\sim 10^{-6}$ A cm^{-2}, while on Pt it is approximately 4 orders of magnitude higher. Let us examine how this would influence the corrosion of iron in an acidic solution when it is electrically coupled to platinum. This is illustrated in Figure 17.11. Before the electrical coupling of Fe to Pt, only Fe dissolves and hydrogen bubbling is noticed around it; nothing is evident around Pt. The reversible potential for the HER depends on pH and can be calculated using Eq. (17.50). At pH = 0, for example, $E_{rev} = 0$ for this reaction. The reversible potential of the iron dissolution reaction can be calculated using Eq. (17.77). The corrosion potential and corrosion current density will be established at the intersection of the iron oxidation line and the line of hydrogen reduction on Fe. When we close the switch between Fe and Pt, the HER occurs more intensively on Pt because its j_0 for this reaction is much higher than that of Fe. Note that both the reversible potential of the HER and the slope of the hydrogen reduction line are identical for Fe and Pt. The new corrosion potential and corrosion current density

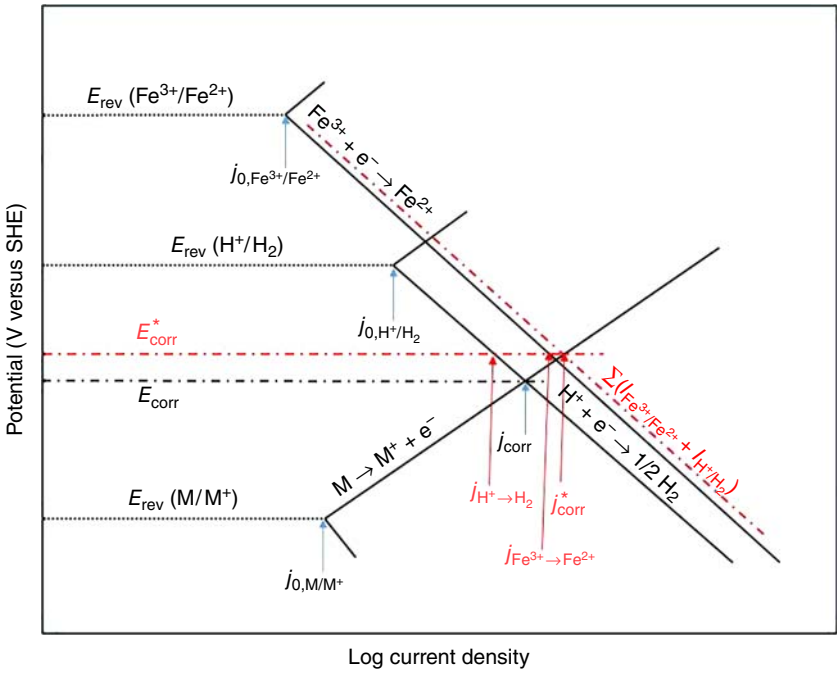


Figure 17.10 Evans diagram showing the effect of adding an oxidizing agent (ferric salt) to the aqueous solution.

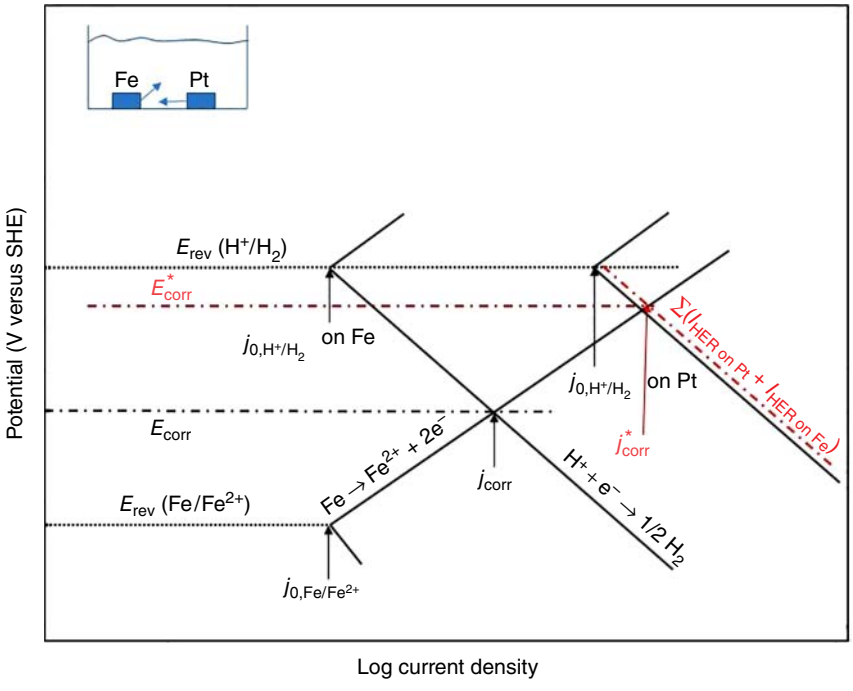


Figure 17.11 Evans diagram illustrating the effect of coupling a corroding metal in a deaerated acidic solution to another metal with a high exchange current density for the HER.

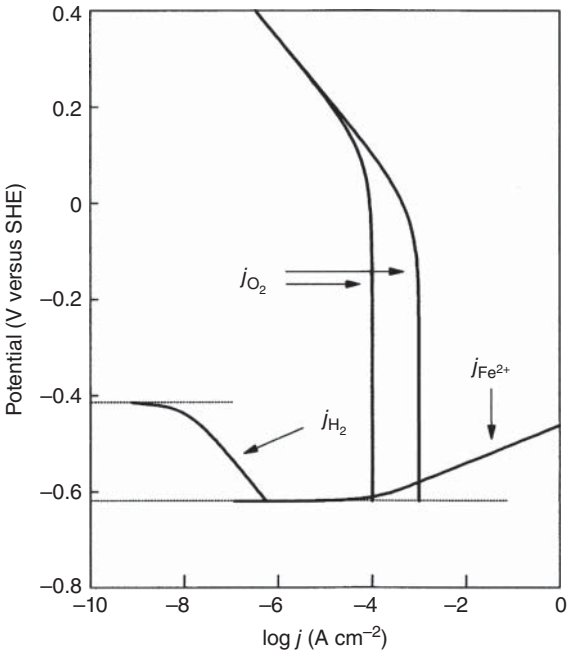


Figure 17.12 Evans diagram for corrosion of iron in neutral, aerated solutions. Two values of the limiting current for oxygen reduction (10^{-4} and 10^{-3} A cm^{-2}) are shown, yielding two different values for E_{corr} and $j_{corr.}$

will be established at the intersection of the iron oxidation line and the sum of the hydrogen reduction lines. Both E_{corr} and j_{corr} increase as a result of this coupling.

In practice, there is no interest in the corrosion of iron in 1 M HCl (neither is it practical to reduce the corrosion rate through the use of Pt). The pH in typical environments is in the range of 5–9. Under these conditions, the rate of hydrogen evolution may be very slow, and oxygen reduction can become the main cathodic process, controlling the rate of corrosion. We recall that the standard potential for oxygen reduction is +1.229 V versus SHE, which leads to a reversible potential of +0.816 V versus SHE at pH = 7. Thus, corrosion occurs at a very high (negative) overpotential with respect to oxygen reduction, and the current is mass-transport limited. Figure 17.12 presents the Evans diagram for iron in neutral, aerated solutions. The exchange current densities for oxygen reduction and hydrogen evolution were taken as $j_0 = 1 \times 10^{-10}$ A cm^{-2} and $j_0 = 1 \times 10^{-8}$ A cm^{-2}, respectively. The contribution of the HER to the measured corrosion current is, therefore, quite negligible in neutral aerated solutions.

The mass-transport-limited current density for oxygen reduction is independent of the kinetic parameters for this reaction. Instead, it depends on factors such as the concentration and the diffusion coefficient of oxygen in the medium. For example, increasing the temperature decreases the solubility of O_2 in water, but increases its diffusion coefficient. It depends also on the rate of flow of the liquid in a pipe or around a sailing ship or a structure immersed in a river. Figure 17.13 illustrates the effect of stirring on the corrosion rate of metal M when the cathodic reaction, oxygen reduction, is under mass-transport limitation. For stirring velocities 1–3, the corrosion current density is simply the limiting current density of the cathodic reaction; it increases as the stirring

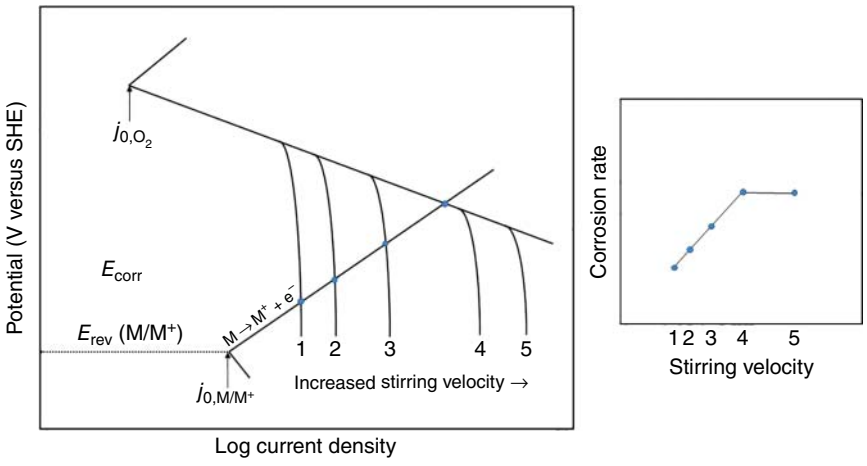


Figure 17.13 Evans diagram for corrosion of a metal in aerated solutions, illustrating the effect of stirring velocity on the limiting current density of oxygen reduction and, consequently, on the corrosion rate.

velocity is increased. At stirring velocity 4 and above, the corrosion rate will no longer be affected by the stirring velocity. It is also evident that we can draw oxidation lines for other metals, whose intersections with the reduction lines 1–3 will yield the same corrosion current density.

Although hydrogen evolution and oxygen reduction are the two reactions most often involved in environmental corrosion, other electroactive materials may take part in the process, enhancing or retarding it, depending on whether they can be reduced or oxidized, respectively, in the range of potential where corrosion takes place.

Next, let us discuss the mixed-potential behavior of galvanic couples. This is illustrated in Figure 17.14 for the bimetallic couple of Fe and Zn in standard conditions. This Evans diagram also explains the higher corrosion rate of iron than zinc in hydrochloric acid solution when uncoupled: despite the more positive reduction potential of iron, the evolution of hydrogen on iron has a significantly higher exchange current density; therefore, $j_{corr,Fe} > j_{corr,Zn}$ in the uncoupled state. As a result of coupling, $E_{corr,couple}$ and $j_{corr,couple}$ are established. The same potential must exist on each constituent of the system. Therefore, we can conclude that as a result of coupling, (i) the corrosion rate of iron decreased from $j_{corr,Fe}$ to $j_{corr,Fe\ couple}$; (ii) the corrosion rate of zinc increased from $j_{corr,Zn}$ to $j_{corr,Zn\ couple}$; (iii) the rate of hydrogen evolution on Zn decreased from $j_{corr,Zn}$ to $j_{corr,H_2\ on\ Zn\ couple}$; and (iv) the rate of hydrogen evolution on Fe increased from $j_{corr,Fe}$ to the point of intersection of line c with the line $E_{corr,couple}$. Thus, Zn acts as sacrificial anode and lowers the corrosion rate of Fe. This is one type of cathodic protection.

We can similarly use the Mixed-Potential Theory also to estimate the impressed current in cathodic protection necessary to decrease the corrosion rate of iron to zero. This means that we should apply cathodic overpotential so that the actual potential will be more negative than E_{rev}, and the metal will be in its immunity domain.

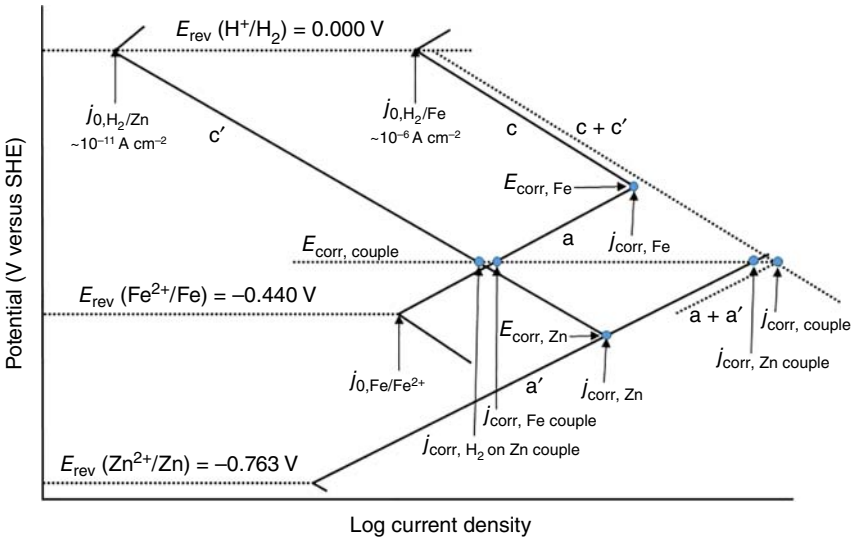


Figure 17.14 Evans diagram showing the mixed potential for the bimetallic galvanic couple of iron and zinc. This figure also explains the higher corrosion rate of iron than zinc in HCl solution. Despite the more positive reduction potential of iron, the evolution of hydrogen on iron has a higher exchange current density.

Finally, let us demonstrate the rationale of using currents, instead of current densities, in Evans diagrams of galvanic couples. In Figure 17.15 we refer to the corrosion of zinc in a deaerated acid solution, with and without coupling to Pt. The exposed surface area of the zinc anode is 1 cm^2. As a result of coupling to Pt of the same surface area, the corrosion current increases from $I_{corr,Zn}$ to $I_{corr,\ Zn:Pt\ 1:1}$, because hydrogen evolution is faster on Pt than on Zn. Since the surface area of the anode did not change, this implies also the same increase in its corrosion current density, or in other words – in its corrosion rate. The exchange current is proportional to the electrode area. Therefore, if we increase the area of Pt to 10 cm^2, the exchange current of the HER on Pt will increase by a factor of 10 too. Consequently, the corrosion current will increase by a factor of 10, to $I_{corr,\ Zn:Pt\ 1:10}$. Since the area of the cathode was increased by a factor of 10, the current density on the cathode will remain the same as before. In contrast, since the area of the anode was not changed, the current density on the anode will increase by a factor of 10, and it will undergo accelerated corrosion. This is an important consideration in the design of systems where dissimilar metals are brought in direct electrical contact – we should make the $A_{cathode}$-to-A_{anode} ratio as small as possible.

17.4.5 Passivation and its Breakdown

Chemical passivation was discovered more than 200 years ago. In 1790 James Keir discovered that treatment of iron with nitric acid produces a peculiar condition in which the metal loses its power of precipitating silver from silver nitrate, although retaining its "metallic splendor." It was Christian Friedrich Schönbein who coined the word "passivity" in 1836, but it was Michael Faraday

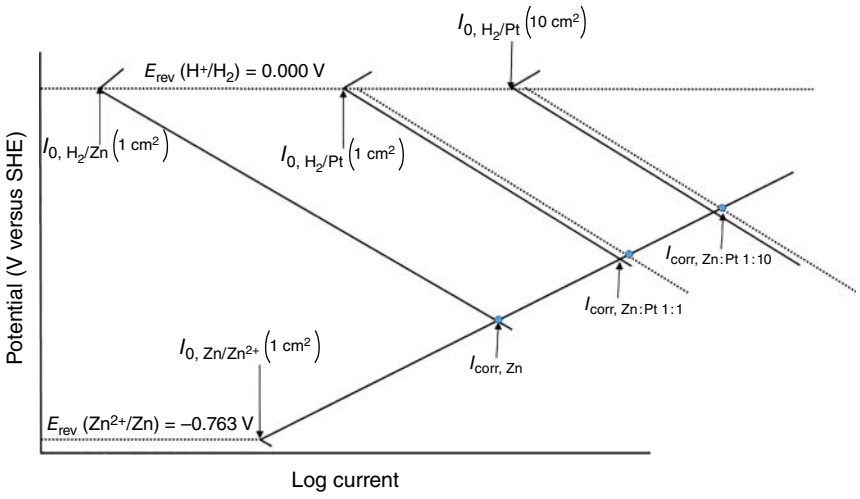


Figure 17.15 Evans diagram demonstrating the effect of the anode-to-cathode surface area ratio of a galvanic couple on the corrosion rate of the anode, as well as the importance of presenting the Evans diagram in terms of currents, and not current densities.

who brought the phenomenon to the forefront of electrochemistry, also in 1836. Faraday established the theory that the passive condition was brought about by a film of oxide, or by a similar state of affairs. The presence of an oxide film, several nanometers thick, on the surface of passivated metals was confirmed experimentally in the 1960s.

In a famous experiment, a piece of iron placed in concentrated nitric acid was found to be passive, while the metal dissolved readily in dilute HNO_3, with copious evolution of hydrogen. This type of behavior can be demonstrated in a very simple, yet quite spectacular, experiment. Nitric acid of various concentrations, from 1 mM to 70%, is introduced into a series of test tubes, and an aluminum wire is placed in each solution. No reaction is observed in the most dilute solutions. As the concentration is increased, however, hydrogen evolution becomes visible. At even higher concentrations, reduction of the acid takes place, in addition to hydrogen evolution. This is evidenced by the liberation of a brown gas, NO_2, which is one of the reduction products. When the concentration has reached 35%, the reaction suddenly stops. There is no gas evolution and the surface of the metal is not attacked. Accurate measurements show no mass loss when aluminum is kept in such solutions for months. Aluminum is passivated in concentrated HNO_3. A thin oxide film is formed on the surface, and further attack is prevented.

Two types of passivity have been defined by Uhlig and Revie: (i) a metal is passive if it substantially resists corrosion in a given environment resulting from marked anodic polarization (low corrosion rate, noble potential); (ii) a metal is passive if it substantially resists corrosion despite a marked thermodynamic tendency to react (low corrosion rate, active potential). Metals or alloys exhibiting type 1 passivity include Ni, Cr, Ti, Fe in oxidizing environments (e.g. chemical passivating solutions such as chromate), stainless steels, and many others. Type 2 examples are Pb in H_2SO_4 and Fe in an inhibited pickling acid.

Passive films formed in aqueous solutions usually consist of an oxide or a mixture of oxides, usually in hydrated form. Passive layers are electronic insulators or semiconductors depending on the band gap of their constituents. Valve metals (Al, Hf, Nb, Ta, Ti, Zr) and even pure Sn may be anodized to potentials much above 1.5 V without oxygen evolution, some such as Al or Ta even to much more than 100 V. Other metals such as Fe, Co, Ni, Cr, Cu, and Ag form semiconducting oxides with a sufficiently small band gap to allow electron transfer reactions with redox systems, including oxygen evolution. Some are even metallic conductors (e.g. IrO_2, RuO_2, and PbO_2). Nickel, chromium, and their alloys with iron (notably, the various kinds of stainless steel) can be readily passivated and, in fact, tend to be spontaneously passivated upon contact with water or moist air. It should be noted that passivation does not occur when chloride ions are introduced into the solution, and a preexisting passive film may be destroyed.

Electrochemical passivation is in many ways similar to chemical passivation. As the potential of an iron sample is increased in the anodic direction, the rate of dissolution increases, reaches a maximum, and then decreases to a very low value. Further increase of the potential has little effect on the current in the passive region until passivity breaks down, whereupon the current rises rapidly with potential. The sequence of events observed on an iron electrode when its potential is swept very slowly in the positive direction is shown schematically in Figure 17.16. The potential at which the anodic dissolution current has its maximum value is called the *primary passivation potential*, E_{pp}. The corresponding maximum anodic current is referred to as the *critical corrosion current*, j_{cc}. In the passive region, which may extend over half a Volt or more, the current j_{pas} is nearly constant.[10] It starts to rise again at the so-called *breakdown* (or critical pitting) *potential*, E_b, above which pitting or crevice corrosion or both will initiate and propagate, along with oxygen evolution and electrochemical dissolution of the passive film – the occurrence of the latter two depends on the conductivity of the oxide. The anodic process taking place at such high potentials usually involve the transformation of the oxide to a higher oxidation state, which is often more soluble. In contrast, on valve metals the oxide continues to grow in thickness as the potential is increased. If there are no aggressive anions in solution, this anodization process can lead to very thick oxide films, up to tens of micrometers, on which oxygen evolution cannot occur.

Cyclic potentiodynamic polarization curves of the type shown in Figure 17.16 are obtained by sweeping the potential, first in the positive direction, very slowly, at a rate of 0.1–1.0 mV s^{-1}. Even so, steady state is not quite reached, and the values of E_{pp} and E_b as well as the shape of the passive region depend to some extent on the sweep rate. In addition, deaeration of the solution with nitrogen gas before and during the test is a common practice. The environment and alloy composition both affect the anodic polarization part. Increasing the chloride (or some other halides) concentration in the environment will shift the curve to the right and will gradually reduce E_b and the whole passivity region. Increasing the temperature will have a similar effect. Increasing the chromium concentration in the alloy will shift the curve to the left and increase E_b and the passivity region.

10 Very often, metastable pits are apparent by transient bursts of anodic current.

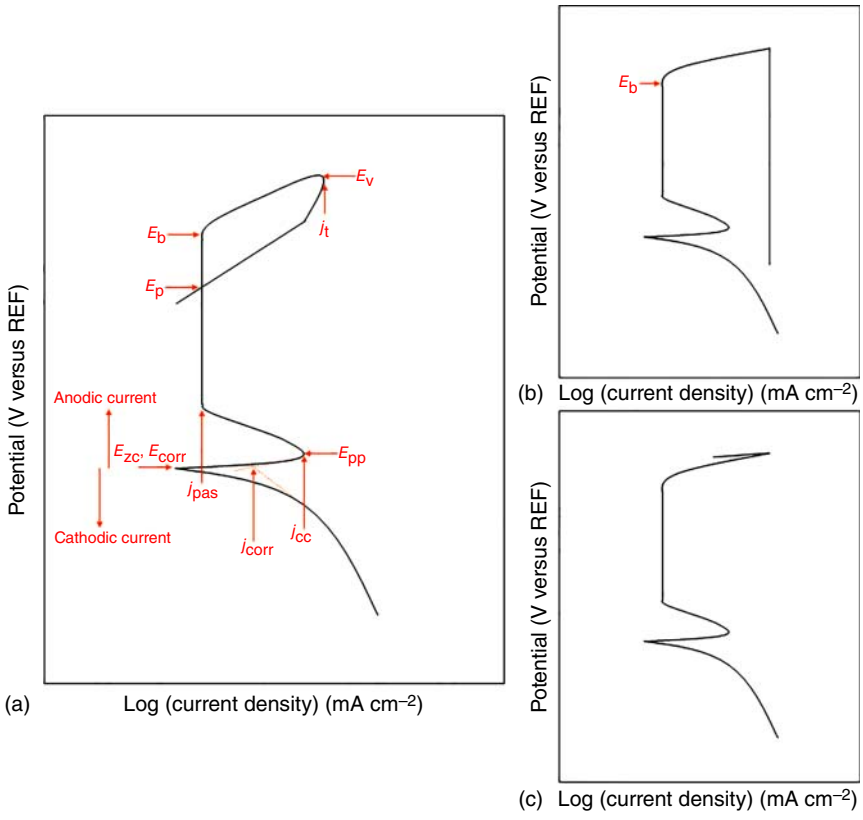


Figure 17.16 Schematics of cyclic potentiodynamic polarization curves of (a) a metal that exhibits a protection potential, (b) a metal that does not exhibit a protection potential, and (c) a metal that repassivates.

Increasing the molybdenum concentration in the alloy will just extend the passivity region to higher potentials, thus improving the resistance of the alloy to localized corrosion.

One of the unique features of most corroding metals undergoing passivation is a region of apparent negative resistance. Looking at Figure 17.16, we note that at potentials anodic to the primary passivation potential E_{pp}, the current density *decreases* with *increasing* anodic potential until it reaches the passive region. This is an unstable region in which the current keeps decreasing with time, even at constant potential, as a result of the formation and growth of the passive film. It should be noted that a classical experimental polarization curve, in which all of the parameters characterizing passivity can be measured as shown in Figure 17.16a, is observed under reducing conditions (e.g. in acids such as HCl), where the predominant cathodic reaction is hydrogen evolution. In contrast, in the presence of oxidizing species (such as dissolved oxygen), some metals (e.g. titanium) and alloys spontaneously passivate and thus exhibit no active region (and no E_{pp}, j_{cc}) in the polarization curve. The oxidizer adds an additional cathodic reaction to the Evans diagram, as we have shown before, and causes the

intersection of the total anodic and total cathodic lines to occur in the passive region (i.e. $E_{corr} > E_{pp}$). The open-circuit dissolution rate under these conditions is j_{pas}, which is often on the order of 0.1 μA cm^{-2} or less.

Another important quantity is the so-called *protection potential*, E_p, sometimes referred to as *repassivation potential*, at which the reverse scan intersects the forward scan at a value that is less noble than E_b. The protection potential cannot be determined if there is no breakdown (see Figure 17.16). The absence of a hysteresis loop indicates repassivation or oxygen evolution. While pitting will occur on a pit-free surface above E_b, it will occur only in the range of potentials between E_p and E_b if the surface is already pitted. An increase in the resistance to pitting corrosion is associated with an increase in E_b. The severity of crevice corrosion susceptibility increases with increasing hysteresis of the polarization curve, i.e. the difference between E_b and E_p. Therefore, a higher value of E_p reflects higher resistance to crevice corrosion.

Controversy concerning the interpretation of cyclic polarization curves has raged for many years. There are those who feel that two distinct potentials do not exist. These workers propose that, when measured correctly, E_b and E_p are one and the same. In standard testing, the nucleation of pits occurs at E_p, but owing to the time necessary for pits to become established, the probability that pits will repassivate, and the finite potential scan rate used, pits do not cause a dramatic increase in the current until E_b. This explanation would rationalize the often-observed effect that increasing the scan rate increases E_b, but not E_p. The classic interpretation is that a material's potential must exceed E_b in order to initiate pits, but if flaws were introduced into the surface in any way, they could propagate at all potentials above E_p. Thus, E_p could be used in design as a protection potential.

It is interesting to consider how changes in the rate of the cathodic reaction can influence the open-circuit corrosion potential and the corresponding corrosion current. This behavior is shown schematically in Figure 17.17. Several cases can be distinguished. Line 1 crosses the anodic dissolution curve in the active region, leading to a substantial rate of corrosion (not good, but predictable). An increase in the rate of the cathodic reaction leads to a shift of the corrosion potential in the positive direction and an increase in the corrosion rate. An example to such a behavior is Ti in HCl or H_2SO_4. Line 2 crosses the line for metal dissolution at three places. Point C is unimportant, since it is unstable, as pointed out earlier. This leaves the system with two stable corrosion potentials. Point B (high j_{corr} in the active region) represents the usual situation found for an actively corroding metal. Increasing the cathodic current from line 1 to 2 raises the corrosion rate from point A to B. But, line 2 also crosses the anodic dissolution curve at point D, in the passive region (low $j_{corr} = j_{pas}$). As a result, a different corrosion potential, corresponding to point D, could also be established. Note that, in the present example, the corrosion potential at point B represents a corrosion rate about 300 times that represented by point D, and in practice the ratio could be substantially higher. Where will the system actually settle? That depends on the initial conditions. If the metal is initially passivated (by oxidizing it chemically or by increasing the potential in the anodic direction), it can remain passivated with the corrosion potential at D. If it is initially in the active region, it can establish

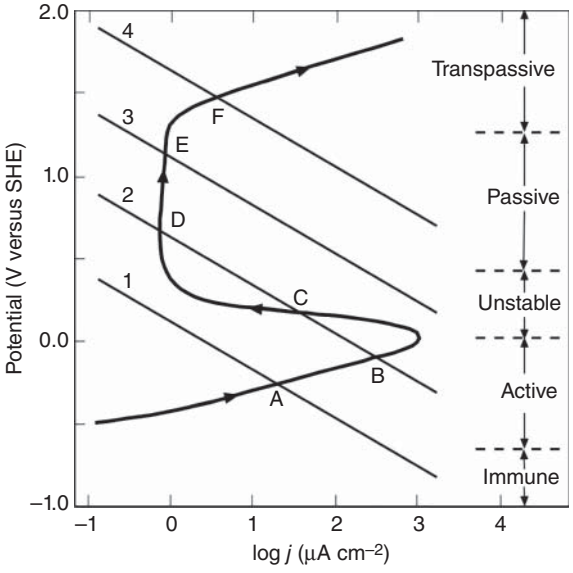


Figure 17.17 Schematic presentation of the effect of changing the rate of the cathodic reaction on the corrosion potential and the corrosion current density in a system undergoing passivation.

its corrosion potential at B. This is referred to in the literature as unstable passivation, because passivation can be lost by transition from the passive region of the potential at point D to an equally stable active corrosion potential at point B. Examples include Cr in air-free H_2SO_4 and Fe in dilute nitric acid. Line 3 represents stable passivation. The anodic and cathodic lines cross at a single point and a corrosion potential is set up at point E, well inside the passive region. This is the most desirable system. Examples include stainless steels and Ti in acid solutions containing oxidizers. However, increasing the cathodic current even more could drive the potential to point F in the transpassive region, where corrosion and pitting could occur.

Many theories of metal passivity have been presented in the literature. The major theories are the oxide film theory and the adsorption theory. According to the *oxide film theory*, the passive film is always a diffusion-barrier layer of reaction products that separate the metal from its environment and slows down the rate of reaction. According to this theory, the following processes are driven by the electric field across the film: (i) entry of metal atoms into the film as cations at the metal/film interface; (ii) transport of the metal cations or of oxygen anions through the oxide; and (iii) dissolution of metal cations from the film at the film/environment interface. On the other hand, the *adsorption theory* holds that chemisorbed films displace the normally adsorbed water molecules and slow down the rate of anodic dissolution involving hydration of metal ions. This theory was supported by Langmuir who showed the oxygen adsorbed on tungsten to exhibit reduced chemical reactivity compared to oxygen in the oxide WO_3. In 1946, Uhlig proposed that an adsorbed oxygen film is the primary source of passivity. Such a film forms preferentially on the transition metals in accordance with their uncoupled d-electrons interacting with oxygen to form a stable bond, combined with their high heats of sublimation favoring retention of metal atoms in their lattice in preference to their removal to form an oxide

lattice. He proposed that a film of adsorbed oxygen atoms markedly decreases the exchange current density in the Tafel equation, and hence increases anodic polarization in accordance with the requirements of passive behavior. For non-transition metals with filled d-levels, such as Cu or Zn, the heats of oxygen adsorption are expected to be lower, and the formation of oxides is less favorable. Such metals do not exhibit thin-film passivity.

17.5 Corrosion Measurements

Corrosion measurements are conducted for different purposes, such as (i) quality control of materials and processes, (ii) determination of lifetime of equipment or product, (iii) failure analysis, (iv) mapping of available materials to select the most appropriate one for a certain application, (v) evaluation of new materials or processes, (vi) evaluation of the service environment and changes in it, (vii) determination of the most economical mean to reduce corrosion, (viii) study of corrosion mechanisms, etc. Corrosion test methods can be divided into electrochemical and non-electrochemical methods.

17.5.1 Non-Electrochemical Tests

Non-electrochemical techniques include direct immersion of samples in the test environment, either in the laboratory or in the plant/field. These samples sometimes have an artificial crevice generated with a serrated washer. They may be welded to determine the effects of welds and their heat-affected zone (HAZ). Real-time information can be obtained using electrical resistance probes. Heat transfer effects can be evaluated by having a test sample that is exposed to the corrosive media on one side, while the other side is heated or cooled. Stressed samples are used to evaluate the resistance to stress corrosion cracking (SCC), hydrogen embrittlement (HE), or corrosion fatigue (CF).

Exposure/immersion tests (ASTM G31) involve the fabrication of coupons, usually in the form of sheets, plates, or U-bend specimens. Corrosion test racks are used to mount the coupons while they are electrically insulated from one another and either fully or partially immersed in the solution, or only exposed to the vapor above the solution. In order to simulate temperatures and pressures above the atmospheric boiling point, tests can be carried out in an autoclave. After exposure, both mass change measurement and visual inspection for surface defects, corrosion products, and color changes are carried out. Corrosion test racks make it easy to evaluate how corrosion would affect different alloys and material finishes under identical conditions.

Atmospheric test racks[11] are also used (ASTM B537, G4). Panel specimens are usually placed in racks at a 30° angle to the horizontal, facing the source of corrosive elements. They are electrically insulated from the racks on which they are mounted and are arranged so that drip from neighboring panels does not contaminate them. Cylindrical specimens are mounted horizontally, facing the same

11 http://corrosion-doctors.org/Corrosion-Atmospheric/Corrosion-tests.htm

direction as the panels. They may be exposed fully to the weather or be partially sheltered, depending upon the requirements of the test.

Corrosion cabinet tests are conducted in closed cabinets where the conditions of exposure are controlled. They are mostly designed to accelerate specific corrosion situations while trying to emulate as closely as possible the corrosion mechanisms at play. Cabinet tests are generally used to determine the corrosion performance of materials intended for use in natural atmospheres. Modern surface analysis techniques can be quite useful to ascertain that the corrosion products have the same morphologies and crystallographic structures that are typically found on equipment used in service. Controlled humidity tests, corrosive gas tests, and salt spray testing are the three major types of corrosion cabinet tests.

In controlled humidity tests (e.g. ASTM A380), the mechanism used to control the humidity moves chamber air via a blower motor and passes it over a heater coil in the bottom of the chamber with an atomizer nozzle fogging into this air stream. There are at least 15 ASTM standards relating different variations of creating and controlling fog and humidity in cabinets for corrosion testing of a broad spectrum of products, from decorative electrodeposited coatings to the evaluation of the corrosivity of solder fluxes for copper tubing systems.

In corrosive gas tests (e.g. ASTM B775, B799, B827, and G87), controlled amounts of corrosive gases are added to humidity to replicate more severe environments. Some of these tests are designed to reveal and amplify certain characteristics of a material.

Salt spray testing is the oldest and most wildly used cabinet test (e.g. ASTM B117, B368, D1654, D2059, D2803, D2933, G85, and G87; ISO 4536, 4623, 6988, and 7384). This test introduces a spray in a closed chamber where some specimens are exposed at specific locations and angles. The concentration of the NaCl solution has ranged from 3.5% to 20%. Although used extensively for specification purposes, results from salt spray testing seldom correlate well with service performance. Hot, humid air is created by bubbling compressed air through a bubble (humidifying) tower containing hot deionized water. Salt solution is typically moved from a reservoir through a filter to the nozzle by a gravity-feed system. When the hot, humid air and the salt solution mix at the nozzle, it is atomized into a corrosive fog. This creates a 100% relative humidity condition in the exposure zone. For a low humidity state in the exposure zone of the chamber, air is forced into the exposure zone via a blower motor that directs air over the energized chamber heaters.

Non-electrochemical, mechanical methods include constant strain, constant load, and pre-cracked specimens. These specimens are then exposed *in situ* to the environment of study. Standard tensile V-notch specimens (ASTM F519) are often used to determine susceptibility to HE, e.g. after electroplating. When tested under constant load, usually 75% of the fracture load, for at least 200 hours – no cracking should be observed. The most common C-ring sample for determination of susceptibility to SCC (ASTM G38, G44) is a constant-strain specimen with tensile stress produced on the exterior of the ring by tightening a bolt centered on the diameter of the ring. A stress of 75% of the ultimate tensile strength (UTS) is often preset, and the sample is subsequently immersed in

the solution, removed periodically, and inspected for visual evidence of stress corrosion cracks. The alternate immersion test usually utilizes a 1-hour cycle that includes a 10-minutes period in solution and a 50-minutes period out of solution.[12]

17.5.2 Electrochemical Tests

The most common electrochemical corrosion tests include the OCP measurement, potentiodynamic polarization scans (Tafel extrapolation, cyclic scans for determination of suscepitibility to localized corrosion, etc.), linear polarization resistance (LPR), corrosion current monitoring, controlled potential tests for cathodic and anodic protection, electrochemical impedance spectroscopy (EIS), electrochemical quartz crystal microbalance (EQCM), rotating cylinder electrode (RCylE) for the study of velocity effects, electrochemical noise (EN) measurement, galvanic probes (zero resistance ammeter, ZRA), and harmonic distortion analysis (HDA). Some fundamental measurements in electrochemistry were already described in Chapter 3, including the three-electrode setup, the residual potential drop jR_S, and reference electrodes; the meaning of the SHE in Section 2.3; potentials of different reference electrodes in aqueous solutions at 25 °C on the inner cover; potential control versus current control in Section 4.2.1; the empirical Tafel equation in Section 5.2.1; the limiting cases of the Butler–Volmer equation (namely, Tafel extrapolation and LPR) in Sections 5.2.4 and 17.4.2; linear potential sweep and cyclic voltammetry in Chapter 14; EIS in Chapter 15; and EQCM in Chapter 16. Hence, only a complementary description of selected tests employed in corrosion studies is provided in this section.

17.5.2.1 Open-Circuit Potential (OCP) Measurements

OCP is the potential of the working electrode relative to the reference electrode when no potential or current is being applied to the cell. In the case of a reversible electrode system, the OCP is also referred to as the equilibrium potential. Otherwise, it is called the rest potential (E_r) or the corrosion potential, depending on the studied system. Metals with nobler (more positive) OCP are more thermodynamically stable than those with more negative OCP; the former will be less susceptible to uniform corrosion. The first step in most electrochemical tests is to measure the OCP. It is very important to allow sufficient time for the OCP to stabilize before beginning the next experiment. This could take between minutes and days. A stable OCP is taken to indicate that the system being studied has reached "steady state," i.e. the various corrosion reactions have assumed a constant rate.

Figure 17.18 shows, as an example, the time dependence of the OCP of Ti–5Ag and Ti–5Ag–35Sn (wt%) alloys prepared by three-dimensional printing followed by liquid-phase sintering or liquid tin infiltration. The OCP of pure Ti control sample is also shown, for comparison. The test solution was non-deaerated saline (0.9% NaCl, pH 5.5) at 37 °C. From Figure 17.18, it is evident that the OCP of both pure Ti and Ti–5Ag alloys increases over time (i.e. they become more noble due

12 N. Eliaz et al. *Eng. Failure Anal.* **9** (2002) 167.

to formation of a stable passivation layer). Because the oxygen reduction rate on titanium is much lower compared to any other oxidizer, about a week is required for the reduction reaction to attain steady state. The addition of Ag affects the steady-state passivation behavior. Over the whole 20-day period, the Ti–5Ag alloy sintered at 1300 °C shows the less negative OCP among all five materials. In principal, the Ti–Ag system could be regarded as galvanic coupling between a passive metal and a noble metal. Typically, coupling of Ti with noble metals such as Pt, Au, or Pd results in spontaneous passivation, since these metals act as good catalysts (i.e. have higher j_0) for the hydrogen and oxygen reduction (cathodic) reactions. The inclusion of 5 wt% Ag ($E^0 = +0.799$ V versus SHE, see Table 17.1) is thus sufficient to increase the corrosion potential of Ti–Ag. Because sintering at 1300 °C results in a more chemically homogeneous material with lower porosity compared to sintering at 1150 °C, the effect of Ag on the corrosion potential is more pronounced for the former alloy. In the case of pure Ti, as the surface oxide layer becomes thicker, the kinetics of mass transport through the layer becomes slower, thus resulting in a decrease in the passivation current density. On the other hand, Ag enhances the oxygen anion uptake in the n-type oxide layer on the Ti–Ag alloy, thus consuming oxygen anion vacancies. This vacancy deficit could result in a reduced passivation current density. If the solution was aerated, the combination of a more significant contribution of the hydrogen reduction reaction and corrosion potential increase due to the presence of Ag could have stabilized the Ti–Ag alloy within the passivity region. Thus, the positive effect of Ag addition could have been even more distinct. The Ti–Ag–Sn alloys also exhibit a galvanic couple behavior, but it is much different from that observed in the Ti–Ag system. At a constant potential, the cathodic polarization curve of Sn typically exhibits a lower current density, while the anodic polarization curve of Sn exhibits a much higher current density, compared to Ti. Therefore, the corrosion potential is likely determined by Sn anodic reaction and Ti cathodic reaction. Consequently, the Ti–Ag–Sn alloys do not exhibit any passivation before Sn dissolution is significantly increased.

17.5.2.2 Polarization Tests

The understanding gained by considering the Evans diagrams allows us to measure the corrosion current directly. First, we must realize that the corrosion potential is in fact the OCP of a system undergoing corrosion. It represents steady state, but not equilibrium. It resembles the reversible potential in that it can be very stable. Following a small perturbation, the system will return to the open-circuit corrosion potential, just as it returns to the reversible potential. It differs from the equilibrium potential in that it does not follow the Nernst equation for any redox couple, and there is both a net oxidation of one species and a net reduction of another.

Consider now the current–potential behavior of a system close to E_{corr}. Assuming that the two partial currents are in their respective linear Tafel region, one can write

$$j_{\text{c}} = j_0 \exp\left[-\frac{E - E_{\text{rev}}(c)}{b_{\text{c}}}\right] \tag{17.108}$$

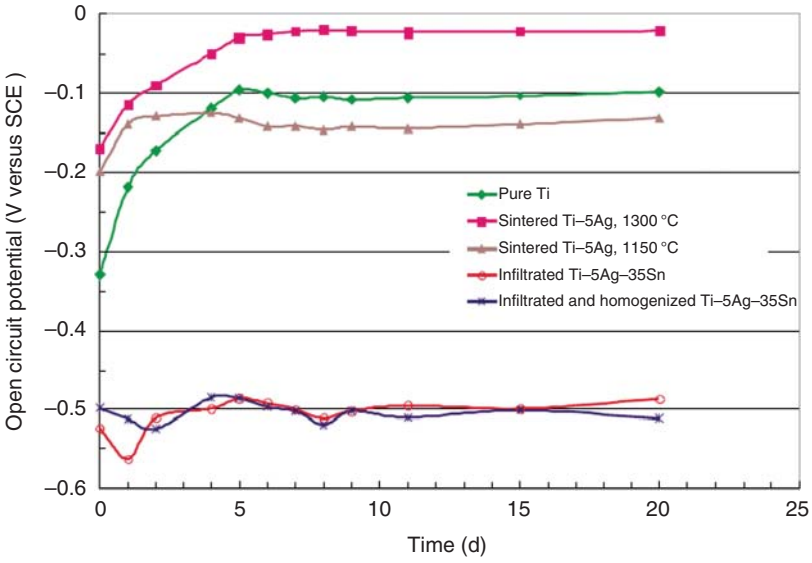


Figure 17.18 The time dependence of the OCP of three-dimensional printed Ti-based alloys in saline solution at pH 5.5, $T = 37\,^{\circ}\mathrm{C}$. Source: Reprinted with permission from Hong et al., *Corros. Sci.* **43** (2001) 1781.

The potential E near the corrosion potential can be written as

$$E = E_{\mathrm{corr}} + \Delta E \tag{17.109}$$

Substituting in Eq. (17.108), one has

$$j_{\mathrm{c}} = j_0 \exp\left[-\frac{E_{\mathrm{corr}} - E_{\mathrm{rev}}(c)}{b_{\mathrm{c}}}\right] \cdot \exp\left(-\frac{\Delta E}{b_{\mathrm{c}}}\right) \tag{17.110}$$

The cathodic current density at the corrosion potential is equal to the corrosion current density

$$j_{\mathrm{corr}} = j_0 \exp\left[-\frac{E_{\mathrm{corr}} - E_{\mathrm{rev}}(\mathrm{c})}{b_{\mathrm{c}}}\right] \tag{17.111}$$

Hence, Eq. (17.110) can be written in the simple form

$$j_{\mathrm{c}} = j_{\mathrm{corr}} \exp\left(-\frac{\Delta E}{b_{\mathrm{c}}}\right) \tag{17.112}$$

This is very similar to the Tafel equation, written for a cathodic process as

$$j_{\mathrm{c}} = j_0 \exp\left[-\frac{\eta}{b_{\mathrm{c}}}\right] \tag{17.113}$$

The corrosion current density, like the exchange current density, is an internal current, which is not observed in the external circuit. The potential difference ΔE is the difference between the applied potential and the OCP, just as η is the difference between the applied potential and the reversible potential. The big difference is that j_{corr} is equal to the anodic and cathodic currents of *two entirely*

different processes, whereas j_0 represents the equal anodic and cathodic currents of the *same reaction* at the equilibrium potential.

Following the same arguments, we can derive for the anodic current density an expression equivalent to Eq. (17.112), namely

$$j_a = j_{corr} \exp\left(\frac{\Delta E}{b_a}\right) \tag{17.114}$$

The net current density observed at a potential E close to the open-circuit corrosion potential is, hence,

$$j = j_a - j_c = j_{corr}\left[\exp\left(\frac{\Delta E}{b_a}\right) - \exp\left(-\frac{\Delta E}{b_c}\right)\right] \tag{17.115}$$

If the cathodic reaction is mass-transport-controlled (i.e. $b_c \to \infty$), we can derive a similar expression for the micro-polarization region. For the anodic current Eq. (17.114) holds, and for the cathodic reaction one has $j_c = j_L = j_{corr}$. The total current is, hence

$$j = j_a - j_c = j_{corr} \exp\left(\frac{\Delta E}{b_a}\right) - j_{corr} = j_{corr}\left[\exp\left(\frac{\Delta E}{b_a}\right) - 1\right] \tag{17.116}$$

When we linearize the exponent, this gives rise to

$$\frac{j}{j_{corr}} = \frac{\Delta E}{b_a} \quad \text{or} \quad j_{corr} = \frac{b_a}{R_p} \tag{17.117}$$

Experimental studies usually yield good agreement between the rates of corrosion obtained from polarization resistance measurements and those derived from mass-loss data, considering that the Tafel slopes for the anodic and the cathodic processes may not be known very accurately. It cannot be overemphasized, however, that both methods yield *the average rate of corrosion* of the sample, which may not be the most critical aspect when localized corrosion occurs. In particular, it should be noted that at the open-circuit corrosion potential, the total anodic and cathodic *currents* must be equal, while the local *current densities* on the surface can be quite different. This could be a serious problem when most of the surface acts as the cathode, and small spots (e.g. pits or crevices) act as the anodic regions. The rate of anodic dissolution inside a pit can, under these circumstances, be hundreds of times faster than the average corrosion rate obtained from micro-polarization or mass-loss measurements.

The corrosion potential, E_{corr}, is the potential of a corroding surface in an electrolyte, relative to a reference electrode. It is deduced either from the plateau in the potential transient when the working electrode is not polarized, or from Tafel extrapolation of the anodic and cathodic curves in potentiodynamic polarization curves. The current density at the corrosion potential, j_{corr}, is also deduced from potentiodynamic polarization curves and is directly proportional to the corrosion rate. The higher E_{corr} and the lower j_{corr} are, the better the corrosion performance of the material is.

Cyclic potentiodynamic polarization and Tafel extrapolation measurements are described in standards such as ASTM G5, G61, G102, and F2129. In potentiodynamic polarization tests, the potential of the test specimen is usually

controlled, and the corrosion current is measured by a potentiostat. The potential is usually first scanned in the positive (forward) direction until a predetermined potential (or current density), usually within the transpassive region, is reached. Then, in the case of cyclic polarization, the scan is reversed until the specimen repassivates or the potential reaches a preset value. The parameters defined in these experiments were identified in Figure 17.16. The vertex potential, E_v, is a preset potential at which the scan direction is reversed. Similarly, the threshold current density, j_{th}, is a preset current density at which the scan direction is reversed. It should be noted that the scan rate (typically, 0.167 or 1 mV s^{-1}) may affect the E_b value and the shape of the passive region. To avoid intensive hydrogen absorption during the cathodic portion of the curve, in particular on hydride-forming metals such as titanium, it is required to start the polarization only 100 mV below E_r.

Figure 17.19 shows the potentiodynamic polarization curve that was measured following the OCP transient that is shown in Figure 17.18. From this figure it is evident that the curves for the pure titanium control sample and for the Ti–Ag alloy sintered at 1300 °C are similar in shape, although the Ti–Ag alloy exhibits a less active E_{corr}. The Ti–Ag alloy sintered at 1150 °C, however, exhibits a less stable passivation (or metastable pitting) at high potentials. The Ti–Ag alloys sintered at either 1300 or 1150 °C exhibit higher j_{corr} values compared to the titanium control samples, although of the same order of magnitude. From this figure, it is also evident that the corrosion potential is much more active, and the corrosion current density is higher, for the Ti–Ag–Sn alloys, either in the as-infiltrated or in the infiltrated and homogenized conditions.

Cyclic galvanostatic polarization (ASTM G100) and potentiostatic polarization (ASTM F746 and G150) techniques are used to study pitting corrosion. They, however, will not be described here in more detail.

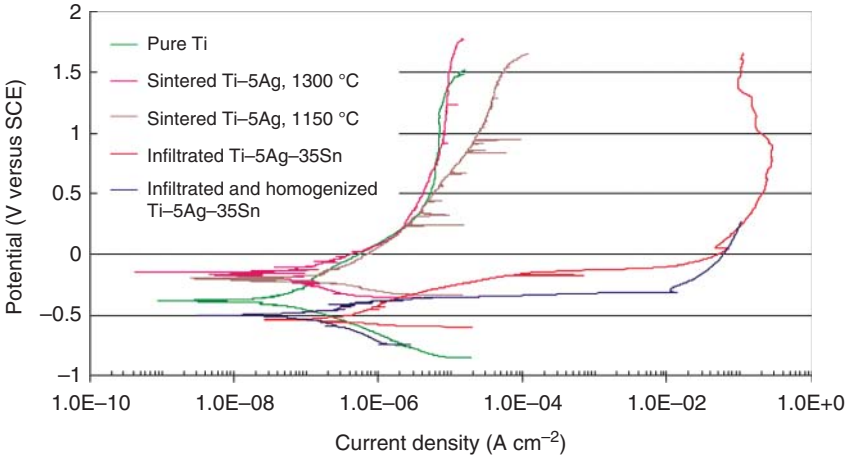


Figure 17.19 Potentiodynamic polarization curves of three-dimensional printed Ti-based alloys in saline solution at pH 5.5, $T = 37$ °C. Source: Reprinted with permission from Hong et al., *Corros. Sci.* **43** (2001) 1781.

17.5.2.3 Linear Polarization Resistance (LPR)

As shown in Sections 5.2.4 and 17.4.2, at low overpotentials a linear approximation results from the Butler–Volmer equation. To experimentally determine R_p, LPR measurement can be conducted. The working electrode is polarized, typically at a scan rate of either 0.166 or 1 mV s^{-1}, within ±20 mV relative to its OCP. The slope of the potential–current density curve at $\Delta E = 0$ is then calculated and is equal to R_p (see Figure 5.1). Subsequently, R_p is used to calculate the corrosion rate using the Stern–Geary equation. To determine R_p, the free corrosion potential (E_{corr} or OCP) must first be empirically determined. Polarization resistance is most commonly run in a three-electrode setup. The WE lead, the WE sense lead, the RE lead, and the CE lead of the potentiostat are always used. If a Faraday cage is used to minimize noise, the floating ground lead will be connected to the cage. The CE sense is turned off during this experiment. It can be left unattached or attached to the cell without issue. More technical information and calculation procedures are available in ASTM G3, G5, G59, and G102 standards.

The advantages of the LPR test include the following: (i) it allows online, real-time measuring of the rate of uniform corrosion, and provides data within a few minutes (in comparison to days or longer in the case of mass loss measurements). (ii) It is very sensitive. (iii) Usually, acceleration factors (such as high temperature) are not necessary. (iv) It is a nondestructive test (NDT) (i.e. measurements may be repeated on the same electrode). (v) It requires simple, easily available equipment.

The limitations of the LPR test include the following: (i) the extent of linearity decreases when the Tafel slope values are low. Deviation from linearity appears at significantly lower overpotentials when the absolute values of the Tafel slopes are lower. Nonidentical Tafel slopes further narrow the linearity region and yield asymmetry around the origin of the curve. (ii) Ohmic errors are present. (iii) It is impossible to measure localized corrosion. (iv) Scales and other deposits formed on the electrode surface give rise to an increase in the electrical resistance irrespective of the change in the corrosion rate of the metal. (v) Redox reactions not involved in the corrosion process per se contribute to the increase of corrosion current, and can result in errors. (vi) The SGC may change over time, and this might result in wrong results. (vii) The test requires prior knowledge of the Tafel constants for calculation of the corrosion rate. (viii) This test is not suitable for electrolyte solutions of low conductivity.

The LPR test is more common than Tafel extrapolation for many industrial applications. Such applications include the monitoring of internal corrosion of subsea production flow lines, corrosion rate of steel reinforcement in concrete, nuclear waste tanks, evaluation of the efficacy of corrosion inhibitors, etc.

17.5.2.4 Zero-Resistance Ammetry (ZRA)

A common current meter (ammeter) measures the current as a voltage drop across an internal shunt resistor. When measuring the current in common electric circuits, this voltage drop can be disregarded. However, it is too high to allow accurate measurement of small currents in galvanic corrosion or local oxide breakdown. An ammeter for such electrochemical measurements should (ideally) have an internal zero resistance. ZRA is a current-to-voltage converter

that produces a voltage output proportional to the current flowing between its input terminals while imposing a "zero" voltage drop (zero resistance) to the external circuit. Any potentiostat can be used as a ZRA: simply connect electrode WE1 to the electrode terminal, WE2 to both the RE plug and the CE plug, set the potential to 0 mV, and switch to potentiostatic mode (counter electrode on). In this setup, a current is displayed on the potentiostat's current meter, and a corresponding voltage is fed to the current output terminal. The voltage between WE1 and WE2 is zero! Thus, we measure the current between two electrodes that are kept at the same potential.

In zero-resistance ammetry (see ASTM G71 and G82), the galvanic current I_{couple} between two dissimilar electrodes, or between two presumably identical electrodes exposed to different environments, is measured by a ZRA. The dissimilar electrodes may represent different chemical compositions, heat treatments, surface conditions, residual stress levels, etc. This technique has been found useful in studying depolarization of the cathode in a galvanic couple and detecting either low levels of dissolved oxygen or the presence of bacteria that increase the coupling current.

17.5.2.5 Electrochemical Noise (EN) Measurements

The term EN is a generic term used to describe the low-level spontaneous fluctuations of potential or current that are easily identifiable in corrosion processes. These fluctuations are related to local variations in the rates of anodic and cathodic reactions as a result of both stochastic processes (e.g. passive film breakdown and repassivation) and deterministic processes (e.g. pit propagation). These fluctuations are usually in the range of microvolts to millivolts, or microamperes to milliamperes, and are of low amplitude and frequency (1 Hz or less). The use of EN measurements for corrosion monitoring has increased since Iverson's first publication in this area (*J. Electrochem. Soc.*, **115**, 617, 1968). A standard test method for EN acquisition was proposed in 2009 (ASTM G199).

The origin of EN may be charge transfer effects (e.g. thermal agitation of charge carriers, or charge being transferred in discrete amounts). In this case, the amount of noise present in a given bandwidth (namely, spectral density) is essentially constant over a wide range of frequencies, and is of low amplitude. A second source of EN relates to surface processes occurring on the electrode, and specifically to their inhomogeneity. These give rise to fluctuations of approximately 1 Hz and below. The observed spectral density of these fluctuations typically varies with frequency, and the amplitude can be much higher than that caused by charge carrier effects. Other sources of EN include metastable pit nucleation and propagation (giving rise to current transients lasting for approximately one second, and involving charge of the order of 10^{-6} C, corresponding to around 10^{12} atoms), fluctuations in the thickness of the diffusion layer (e.g. under turbulent conditions), Brownian motion, hydrogen bubble nucleation, growth and detachment, etc.

Two techniques are used for measuring the electrochemical potential and current noise. In the first technique, two "identical" electrodes are connected through a ZRA. Changes of electric potential and current between these two electrodes are measured over time. It is explicit that it is impossible to create

two ideally "identical" electrodes. Therefore, a second technique was developed, in which one working electrode is connected, together with a counter electrode (usually made of Pt) and a reference electrode, to a potentiostat. This technique is called the potentiostatic EN technique; both potential and current noises are measured on the same working electrode at OCP. Current noise relates to current variations between the working and counter electrodes, whereas potential noise refers to the variations in potential between a working electrode and a reference electrode. Various software tools have been developed for mathematical treatment of the measured values of potential and current noise. While the measurement of EN is relatively straightforward, the data analysis can be complex and inconclusive.

The EN technique differs in many ways from other electrochemical techniques used in corrosion. One important difference is that EN measurements do not require that the working electrode be polarized in order to generate a signal. However, it is also possible to measure current noise under an applied potential, or measure potential noise under an applied current. For simultaneous measurement of electrochemical potential and current noise, a three-electrode setup is required. In field corrosion monitoring, the three-sensor elements are usually made of the same material.

EN is often regarded as the most sensitive technique available for real-time online corrosion monitoring. The "signatures" from EN sensors can be used to derive not only corrosion rates, but also mechanistic information on the corrosion type (general versus localized). Localized corrosion processes, which are hard to monitor with other techniques, give particularly strong electrochemical noise signals. The added advantage of this technique is that it does not polarize the corroding sample and, hence, can be applied to real structures. It is also not affected by conductive deposits. The instrumentation needed to make such measurements is simple.

17.5.2.6 Electrochemical Hydrogen Permeation Tests

Hydrogen is one of the most damaging species in metals and their alloys. When hydronium ions or water molecules are discharged, e.g. during electroplating or corrosion processes, atomic hydrogen is adsorbed at the surface of a metal electrode. This adsorbed atomic hydrogen can either recombine and evolve as molecular hydrogen through the liquid phase, or be absorbed in the metal. The efficiency of hydrogen entry is characterized by the concentration of hydrogen dissolved in the metal just beneath its surface, C_0. Because the chemical potential of atomic hydrogen and its activity are high compared to molecular hydrogen, a very high hydrogen fugacity, f, can be developed at the surface of the metal ($\sim 10^6$ atm has been estimated, which is equivalent to a pressure of 10^4 atm). Consequently, the metal can be charged with hydrogen to concentrations higher compared to charging from a gaseous hydrogen environment. The concentration of hydrogen absorbed in the metal (either occupied in interstitial sites in the lattice or trapped), which is in equilibrium with gaseous hydrogen, can be described by Sievert's law:

$$x = \sqrt{\frac{p}{p_0}}\, e^{\Delta S_s/k}\, e^{-\Delta H_s/kT} \tag{17.118}$$

where x is the solubility of atomic hydrogen in the metal (in terms of atomic ratio H/M), p is the partial pressure of hydrogen gas, ΔS_s is the solution entropy at a standard pressure of hydrogen gas, p_0, ΔH_s is the enthalpy of solution, and k is the Boltzmann constant. The fugacity of hydrogen at the electrode surface during cathodic charging can be calculated using the Nernst equation. However, at high overpotentials (high pressures), the use of the Nernst equation is not valid owing to recombination processes. At a sufficiently high overpotential, the mechanism of hydrogen evolution changes and surface coverage is no longer dependent on overpotential. Consequently, the hydrogen pressure will not increase further above this overpotential.

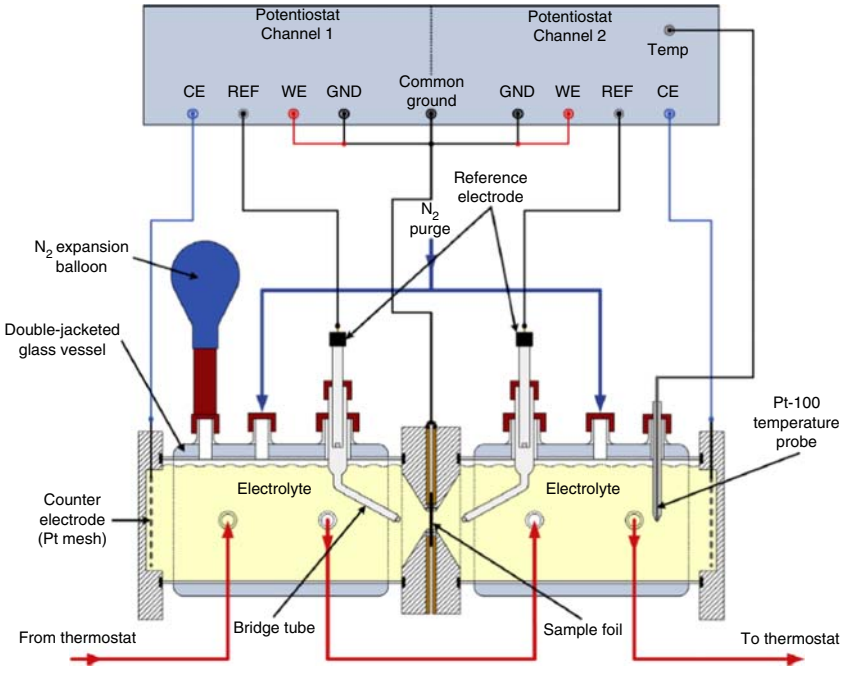
In general, one may distinguish between the following processes in the metal–hydrogen system: (i) the entry of hydrogen from the surrounding environment into the metal, (ii) the transport (diffusion) of atomic hydrogen between interstitial sites in the metal lattice (i.e. interstitial solid solution), and (iii) the trapping of hydrogen at structural defects and/or formation of hydride phases. Only hydrogen in interstitial solid solution moves through the metal according to Fick's laws of diffusion; hence, it is called "diffusible hydrogen." Hydrogen can also be reversibly trapped (e.g. at grain boundaries and around dislocations and nonmetallic inclusions) or form hydride phases. This reversibly trapped hydrogen is more tightly bound within the metal and shows lesser mobility.

Once inside the metal, hydrogen might embrittle it (see Section 17.6.6.1 on HE). Therefore, the characterization of hydrogen in metals is important for the understanding, prediction, and prevention of HE. Reliable information on the above three processes in the metal–hydrogen system can be obtained from electrochemical hydrogen permeation tests. The latter are commonly used to measure the diffusivity (D_H), the solubility (C_H), and the subsurface concentration (C_0) of hydrogen in steels and other materials.

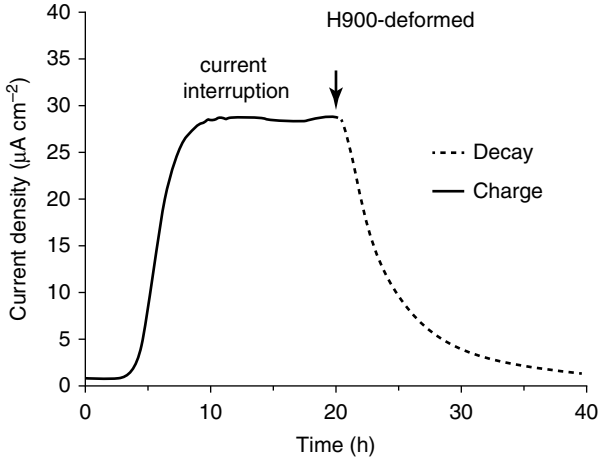
For the electrochemical hydrogen permeation test, the electrochemical cell first proposed by Devanathan and Stachurski (*Proc. R. Soc. London A*, 1962, **270**, 90) is still the most common setup. The cell is comprised of two independent electrochemical cells separated by a thin metal membrane (the specimen). Because there is no pressure difference between both cells, very thin membranes can be used. This reduces the diffusion timescale and increases the permeation flux.

Hydrogen atoms are generated and enter the membrane on its entry side ($x = 0$), diffuse through the membrane, and finally leave the membrane on its parallel exit side ($x = L$), where they are immediately oxidized under potentiostatic conditions. ASTM G148-97(2011) and BS EN ISO 17081:2014E standards describe the setup, procedure, and data analysis in details. Figure 17.20a shows a schematic of the setup.

The proper specimen thickness range should be determined empirically. If the membrane thickness is too high, the experiment duration might be too long. On the other hand, the membrane thickness should be at least five grains large. In a case where this requirement is not complied with, hydrogen may diffuse through grain boundaries, bypassing transport through the grains. In order to avoid edge effects that cause significant hydrogen leakage laterally, the L/r ratio (where L is the membrane thickness and r is the radius of the surface area exposed to the catholyte) should be less than 0.05. For each sample type, it is recommended that



(a)



(b)

Figure 17.20 (a) Schematics of the electrochemical hydrogen permeation cell assembly. The specimen membrane serves as the grounded common electrode of both cells. Source: Reproduced with permission from A. Manhard, S. Kapser, and L. Gao, *J. Nucl. Mater*. **463** (2015) 1057. (b) Electrochemical hydrogen permeation rise and decay transients for Custom 465® martensitic precipitation-hardened stainless steel. Source: Reproduced with permission from S. Ifergane et al., *Electrochim. Acta* **178** (2015) 494.

at least three specimens are tested in various thicknesses, in order to avoid surface and undesired size effects.

The analysis of electrochemical permeation data is very accurate as long as the boundary conditions are perfectly known. To this aim, the membrane should be coated, either only on its exit side or on both sides, with a thin (several nanometers thick) palladium (Pd) coating prior to the permeation test. On the exit side of a steel membrane, Pd reduces the noise levels due to oxidation currents of the steel, ensures high catalytic activity for the hydrogen oxidation reaction, and eliminates the barrier effect of the native oxide film during the discharge cycle. On the entry side, Pd ensures well-defined hydrogen activity and eliminates the barrier effect of the native oxide film. The entry side may be left uncoated if the focus is on the hydrogen uptake, such as under corrosion conditions. Pd is chemically stable in the solution electrolytes that are typically used in permeation tests. Owing to the high diffusivity and solubility of hydrogen in Pd, the concentration of dissolved hydrogen in this very thin layer is approximately homogeneous and in equilibrium with the hydrogen cation activity in the catholyte at all times. The hydrogen concentration at the Pd/steel interface then adjusts so that the respective activities of hydrogen in Pd and steel become equal again. The anodic polarization at the exit side enforces effectively zero hydrogen concentration in the Pd layer on that side, which leads also to zero hydrogen concentration on the exit side of the steel. As long as the solubility and the diffusivity of hydrogen in Pd (and, therefore, its permeability) are orders of magnitude larger than for hydrogen in the specimen material, the influence of the Pd layer on permeability is considered small.

The permeation cell usually consists of two double-jacketed (for heating) glass vessels. The reduction cell usually contains aqueous solution of NaOH (e.g. 0.05 M) or H_2SO_4. To enhance hydrogen uptake, poisons (a term synonymous with undesirable inhibition) such as arsenic (As), selenium (Se), tellurium (Te), and thiourea (CH_4N_2S) are sometimes added to the catholyte. These species inhibit the recombination of adsorbed atomic hydrogen to desorbed molecular hydrogen through either the Volmer–Heyrovsky or the Volmer–Tafel reaction routes (see Eqs. (7.2)–(7.4)), i.e. they hinder the HER and favor the competing absorption route by piling up adsorbed hydrogen and, thus, increasing θ_H.

Another catholyte that has been used is 2 : 1 by volume glycerol:concentrated (85%) phosphoric acid (H_3PO_4). The mobility of molecular hydrogen in this electrolyte is low due to the high viscosity of the electrolyte, according to the Stokes–Einstein law:

$$D = \frac{kT}{6\pi r\eta} \tag{17.119}$$

where D is the diffusion coefficient of the molecule in the electrolyte, r is the radius of the molecule, and η is the electrolyte dynamic (absolute) viscosity. The reduced mobility of hydrogen gas molecules results in increased hydrogen absorption in the membrane. Furthermore, this electrolyte has been reported to have a much lower background current.

The oxidation cell usually contains aqueous solution of NaOH (e.g. 0.05 M). The potentiostat of this cell must be very sensitive to detect the small amounts of

hydrogen that are oxidized. With modern equipment, currents and voltages for both cells can be supplied and measured by a two-channel potentiostat, where the channels are electronically synchronized. Alternatively, two different potentiostats can be connected to the same membrane, but they need to be floating. "Floating ground" potentiostats are isolated from the earth-ground. This allows them to be used to study grounded electrodes or solutions. The potential on each side of the membrane is applied and measured versus a reference electrode in a separate electrolyte. Two double layers are formed, one on each side of the membrane. Thus, the electrochemical potential on the two sides of the membrane can be different.

To achieve sufficiently high sensitivity, the background current established prior to hydrogen transport should be relatively constant and small compared to that of the hydrogen atom oxidation current. This can be achieved by coating the exit side of the membrane with Pd and maintaining an electrode potential where the metal surface is either passive or immune to corrosion. In addition, Ar gas is often purged prior to cathodic polarization and during the entire experiment to remove oxygen. The chosen oxidation potential should be sufficiently positive to ensure that the kinetics of oxidation of hydrogen atoms is limited by the flux of hydrogen atoms. System validation can be done using Pd membranes, for which data is available in the literature for comparison.

In electrochemical permeation tests, the total oxidation current is monitored as a function of time. Figure 17.20b shows a typical rise (charge) and decay (discharge) current transient. During charging, the anodic current rises until steady state is reached, whereas during discharge the current decays until the background oxidation current measured is practically zero (below 0.1 μA cm^{-2}).

In reasonably pure, defect-free metals with a sufficiently low density of microstructural trap sites, atomic hydrogen transport through the material is controlled by lattice diffusion. On the other hand, defects such as dislocations, grain boundaries, inclusions, precipitates, voids, and cracks may act as trap sites for hydrogen, thus delaying its transport. These traps may be reversible or irreversible depending on the binding energy associated with the particular trap sites compared to the energy associated with hydrogen transport through the metal lattice. A value of 50 kJ mol^{-1} is often referred to as the border between reversible and irreversible traps.

The effective diffusion coefficient D_{eff} can be associated with the experimentally determined $D_{\text{lag\_63\%}}$ using the following time-lag equation, which is a solution of Fick's second law for the specific initial and boundary conditions and assuming a constant diffusion coefficient:

$$D_{\text{lag\_63\%}} = \frac{L^2}{6t_{\text{lag\_63\%}}} \tag{17.120}$$

where $t_{\text{lag\_63\%}}$ is the time during charging transient when the current density j reaches 63% of the final steady-state current density, $j_{\max}$. When steady state is achieved during charging, the hydrogen concentration distribution across the specimen is assumed to be linear:

$$c(0,x) = C_0\left(1 - \frac{x}{L}\right) \tag{17.121}$$

where C_0 is the subsurface concentration of atomic hydrogen in interstitial lattice sites on the charging side of the sample (mol m^{-3}). When only reversible trapping is important and the permeation transient can be represented by Fick's second law, the sum of the subsurface concentration of hydrogen in interstitial lattice sites and in reversible trap sites on the charging side of the sample, C_{0R}, can be calculated from

$$J_{ss} = \frac{D_{eff}}{L} C_{0R} \tag{17.122}$$

J_{ss}, the permeation flux of hydrogen at steady state as measured on the exit (oxidation) side of the membrane (mol m^{-2} s^{-1}) is obtained from the maximal current density j_{max}, using the relation $J_{ss} = j_{max}/nF$, where nF is the charge transferred per mole of H.

To verify the applicability of Fick's second law to the permeation transient, the permeation transient is plotted in the form of normalized flux, $J(t)/J_{ss}$, against the logarithm of the normalized time $\tau = D_{eff}\, t/L^2$. The normalized transient is then compared with that derived from Fick's second law for rising transients:

$$\frac{J(t)}{J_{ss}} = 1 + 2\sum_{n=1}^{\infty} (-1)^n \exp(-n^2\pi^2\tau) \tag{17.123}$$

A summation from $n = 1$ to 6 is sufficiently accurate according to ISO 17081 and ASTM G148; in many cases $n = 3$ is acceptable. Eq. (17.123) is the Fourier solution of Fick's second law for the case where the concentration of hydrogen atoms is constant at all times at the entry side of a membrane, and is zero at the exit side.

If the permeation transient is steeper than predicted from Fick's second law, this indicates that trap occupancy is significant. A permeation transient that is less steep than predicted from Fick's second law is often an indication of unsteady surface conditions. At steady state, all of the irreversible traps are occupied. Therefore, if the first permeation transient takes longer than the second permeation transient, it can be deduced that irreversible traps exist and affect permeation. In the absence of void or blister formation or hydrogen-induced cracking, the second and subsequent transients should show the same permeation behavior provided the C_0 values are similar.

17.5.3 Complementary Surface-Sensitive Analytical Characterization Techniques

Many analytical characterization techniques are useful in corrosion studies. Metallographic microscope in the reflected light mode is used to identify macroscopic features such as grain size, inclusions level, and pit size distribution. Differential interference contrast (DIC) microscopy, also known as Nomarski interference contrast, is used to improve the contrast between regions with different chemistries (e.g. several layers in a paint). Scanning electron microscopy (SEM) is used to characterize fracture surfaces (fractography), identify the failure origin zone, etc. When coupled with energy dispersive spectroscopy (EDS), the local chemical composition can be determined, and lateral elemental

maps can be constructed (e.g. to identify high local concentration of chlorine in pits). When coupled with an electron backscatter diffraction (EBSD) detector, microstructural features such as grain size distribution, phase maps, orientation maps, deformation maps, and boundary misorientation distribution can be analyzed. Scanning tunneling microscopy (STM) and atomic force microscopy (AFM) are used to determine the structure and roughness of surfaces at high resolution. While STM is limited to conductive surfaces, AFM is suitable also to nonconductive ones. However, STM in vacuum provides atomic resolution. Electrochemical AFM allows the monitoring of the initiation and propagation of corrosion *in situ* and in real time. X-ray diffraction (XRD) is used to identify the phase content of corrosion products (usually, inorganic). X-ray photoelectron spectroscopy (XPS) and Auger electron spectroscopy (AES) are used to determine the chemical composition and oxidation state (XPS) of surfaces, including oxides, corrosion products, and deposits. Depth profiles enable determination of oxide thickness and variation in composition across the surface layer. Secondary ion mass spectrometry (SIMS) combines simultaneous detection of ions over a virtually unlimited mass range (from hydrogen to uranium), at high mass resolution (up to 0.02 amu), at sensitivities that may be as high as parts per billion to parts per million, with mapping of the lateral distribution of species and with depth profiling. Static SIMS can suggest molecular structures and observe extremely high mass fragments (e.g. of polymers), which dynamic SIMS cannot do. SIMS has been used, for example, to visualize the trapping sites of hydrogen in high-strength steels. The concentration of hydrogen in metals, the temperature range for hydrogen desorption, and the activation energy for hydrogen detrapping can be determined by thermal desorption spectroscopy (TDS). Hydrogen, oxygen, and carbon analyzers can also be used to determine the total amount of these interstitial elements in a small volume of metal.

Vibrational spectroscopy techniques have been used to identify the nature of corrosion products and to quantify their amounts, as well as to identify adsorbed molecules, study the structure of inhibitors, monitor the degradation of organic coatings and paints, etc. Fourier-transform infrared (FTIR) spectroscopy, for example, has also been used to improve the understanding of microbial corrosion through identification of biofilms and their impact on the metal surface. Raman spectroscopy can usually offer a wider spectral range. In studies of local corrosion, confocal Raman microscopy has been used, providing a spatial resolution down to 10–20 nm, which is significantly better than conventional IR microscopy. Enhancing the electric field by roughening the sample surface, known as surface enhanced Raman spectroscopy (SERS), allows huge enhancement, by a factor of 10^6, of the Raman signal, so that even single molecule detection becomes possible. UV–vis spectrometry provides useful information such as the concentration ratio of specific complexes in aqueous solutions as well as the inhibition characteristics and mechanisms of organic compounds.

Scanning electrochemical microscopy (SECM) is a powerful near-field scanning technique. In this technique, current flows through a microelectrode immersed in an electrolyte solution and situated close to a substrate. The substrate can be a conductor, semiconductor, or insulator. The microelectrode, substrate, reference electrode, and auxiliary electrode (and, sometimes, a second

working electrode) comprise the electrochemical cell. A bipotentiostat allows control of the potential of the microelectrode and/or the substrate versus the reference electrode, as well as the measurement of the current flowing between any of the working electrodes and the counter electrode. The microelectrode displacement and its position relative to the substrate are controlled with a precise 3D microstage. The first SECM was designed in 1989 by Bard et al. Since then, it has become a very powerful technique for probing a great variety of electrochemical processes in fundamental and applied electrochemistry, energy storage, materials science, corrosion science, biosensors research, and biophysics. SECM is a powerful technique for studying corrosion processes because it allows for *in situ* characterization of both topography and localized corrosion activity in micrometer and submicrometer ranges. Local corrosion processes can be quantified in real time with high spatial resolution. Electrochemical data may be collected in three main modes: amperometric, potentiometric, and AC. The amperometric and potentiometric modes have been used to characterize passive layers and organic films on metals, as well as defects and pitting, and to sense the local pH on a corroding metal. The more recently introduced AC mode has proved to be useful for studying the insulating/conducting domains as well as changes in electrolyte composition near a corroding metal surface.

17.6 Forms of Corrosion

17.6.1 Uniform (General) Corrosion

Uniform (General) corrosion can be described as a corrosion reaction that takes place uniformly over the surface of the material, thereby causing a general thinning of the component and, eventually, its failure. This type of corrosion is often easy to detect by simple visual inspection, and the corrosion rate can be calculated using Eqs. (17.101)–(17.103), making it straightforward for life prediction.

The most important parameters that affect uniform corrosion are moisture (or relative humidity), temperature, pH, the presence of contaminants in the environment, and stirring velocity.

The hardness of water is determined by the concentration of multivalent cations such as Ca^{2+} and Mg^{2+}. Hard water contains a significant concentration of dissolved salts of calcium and magnesium, principally as bicarbonates, chlorides, and sulfates. Ferrous iron may also be present. The following concentrations of salts in water are commonly used to distinguish between various hardness levels of water: 0–60 ppm (mg L^{-1}) in soft water, 120–180 ppm in hard water, ≥181 ppm in very hard water. Rainwater and distilled water are soft, whereas seawater is very hard. In Israel, for example, drinking water has a total hardness of 255–465 ppm $CaCO_3$. In areas where the water is hard, home water softeners are used, making use of the properties of natural or artificial zeolite minerals. When using hard water, white-gray adherent scales are observed in home kettles, heat exchangers, and condensers. While such scales protect metal surfaces from corrosion, they reduce significantly the heat conductance and might lead to efficiency drop, decrease in flow rate, and even explosion

of boilers. Sodium carbonate, if present, hydrolyzes to produce free alkali that causes caustic embrittlement and failure of boiler plates. Natural soft water, on the other hand, is often very corrosive to metals, because a protective scale is not formed on metal surfaces. There is a debate on whether ion-exchange-softened water is also corrosive. Some argue that since it has not undergone any change in dissolved oxygen, pH, temperature, conductivity, chloride level, and sulfate level, it is not more corrosive than the original hard water.

The concentration of dissolved oxygen gas is another important parameter. The solubility of oxygen in the electrolyte solution decreases as the temperature is raised or the partial pressure of oxygen is decreased; the released oxygen can attack metal surfaces. Typically, the corrosion rate increases as the concentration of dissolved oxygen (or relative humidity) increases. However, above a certain oxygen concentration (e.g. 25–35 ppm in natural water), the corrosion rate may decrease due to passivation of the metal.

SO_2 is a common industrial contaminant. It might react with steel, forming ferrous sulfate ($FeSO_4$), which is partly oxidized to ferric sulfate, $Fe_2(SO_4)_3$. During exposure to water, ferrous sulfate undergoes hydrolysis, forming H_2SO_4. The hygroscopic character of the sulfate increases the time of exposure of the metal to humidity. Consequently, corrosion is accelerated. For example, carbon steel may lose 200 g m^{-2} weight after 80 days of atmospheric corrosion in the absence of SO_2, compared to 1250 g m^{-2} after 80 days in the presence of SO_2. As another example, a critical relative humidity of 60% has been observed for carbon steel in the presence of 0.01% SO_2, due to formation of the hygroscopic corrosion product $FeSO_4$.

Dissolved halide ions such as Cl^-, F^-, Br^-, and I^- increase both the hydrogen evolution rate and the active dissolution rate of carbon steel and tend to breakdown the passive film and induce pitting attack. When the concentration of chloride ions is high, for example, in seawater, passivity is not attained. Nevertheless, halides have also been reported to inhibit the corrosion of some metals in strong acids. This effect depends on the ionic size and charge, the electrostatic field set up by the negative charge of the anion on adsorption sites, and the nature and concentration of halide ions. Given the ability of halides to be adsorbed on metal surfaces, attempts have been made to explain the difference in the action of the halides in terms of their atomic radii and/or the electronegativity.

The pH of the environment is another important parameter. Corrosion may occur both in acidic and in basic environments. The effect of pH on corrosion rate is complex and depends on the corroding metal as well as other variables. All metals may be divided into five groups according to the effect of pH on their oxidation: (i) the corrosion resistance of Au, Ag, Pt, Pd, Rh, Ru, Hg, Ta, Nb, Os, and Ir is almost not affected by the pH. These metals are corrosion resistant within the pH range of 0–14. (ii) Mo, W, and Re are resistant to acids and neutral solutions, but corrode in alkaline solutions. (iii) Mg, Ti, Hf, V, and Bi corrode faster in acidic solutions, but are corrosion resistant in neutral and alkaline solutions. (iv) Fe, Cr, and Mn corrode fast both at $pH < 4$ and at $pH > 13.5$ at temperatures above 80 °C. They also corrode, but at a lower rate, at near-neutral pH (6–8). However, they become corrosion resistant at $pH = 9–13$. (v) Be, Al, Cu, Zn, Cd, Sn, Pb, Co, Ni, Zr, Ga, and In corrode both in acidic and in alkaline solutions

and are therefore called "amphoteric metals." These metals are resistant only to neutral solutions. The region of pH where these metals are corrosion resistant depends on the metal. Temperature and the presence of certain ions in solution may drastically change the pH region of high corrosion resistance.

The corrosion rate is affected not only by the extent of acidity but also by the acid type. Carbon steel tanks are used for transportation and storage of concentrated sulfuric and nitric acids, although Fe belongs to group (iv) of metals that corrode faster in acidic environments. How can it be? Concentrated sulfuric (>60%) and nitric (>40%) acids favor the formation of passive iron oxide layers on iron and carbon steel surfaces. Such layers protect iron alloys from corrosion. Dilution of these acids would result in the destruction of passive iron oxides and give rise to accelerated corrosion of iron.

Carbon dioxide (CO_2) is an acidic natural gas. It currently constitutes about 0.041% by volume of the atmosphere. Rapid industrial development and the use of motor vehicles, and as a result a constant increase in the burning of fuels, give rise to large quantities of CO_2 in the air and to the so-called hothouse effect. When CO_2 dissolves in water droplets in the atmosphere, weak carbonic acid H_2CO_3 is formed, and the pH of the water condensate may decrease to 4. This is one of the causes of natural acidic rains. These rains cause corrosion of metallic constructions. When CO_2 or carbonates diffuse from the atmosphere into reinforced concrete and react with humidity, the local pH can drop from 12.5 to 8.5. Below pH $\sim$ 10.5, the passive layer on iron starts to degrade. The volume of oxidized iron is approximately 10 times larger than the volume of non-oxidized iron. Consequently, swelling and cracking may develop in the concrete, leading to its failure. This phenomenon can be mitigated using cathodic protection, protective coatings, and inhibitors.

One should distinguish between open and closed systems. An open recirculating water system uses the same water repeatedly to cool the furnace system. Airborne gases, especially oxygen, can be absorbed from the air, causing higher corrosion rates. Cooling by evaporation increases the dissolved solids concentration in the water, raising corrosion and deposition. The relatively higher temperatures increase the corrosion potential. The longer retention time and warmer water tend to increase the tendency for biofilms growth. In contrast, in a closed water system, the water circulates in a closed cycle and is subject to alternate cooling without air contact. This can either be accomplished through a closed-circuit evaporative tower or with air-cooled heat exchangers. Air is prevented from being reintroduced to the water, thus reducing corrosion and improving thermal efficiency. Closed systems are also less susceptible to biofouling from slime and algae deposits than are open systems. With the small amount of makeup water required, adequate treatment can virtually eliminate corrosion and the accumulation of corrosion products. Untreated systems can suffer serious corrosion damage due to pitting, crevice, or galvanic corrosion. The three most reliable corrosion inhibitors for closed cooling water systems are chromate, molybdate, and nitrite materials. The level of concentration of these various inhibitors is a direct function of the system component materials and the typical maximum operating temperature of the cooling water.

Temperature is another important parameter. Increasing the temperature will reduce oxygen solubility. In an open system, where oxygen can be released from the system, corrosion will increase up to a maximum at 80 °C, where oxygen solubility is 3 mg L^{-1}. Beyond this temperature, the corrosion rate decreases because the reduced oxygen concentration limits the cathodic reaction. For closed systems, in which oxygen cannot escape, corrosion continues to increase linearly with temperature. Temperature has additional effects. Increasing the temperature will increase the rate of oxygen transport to the metal surface, thus increasing the corrosion rate since more oxygen is available for the reduction reaction. The viscosity of water will decrease with increasing temperature, which will aid oxygen diffusion. Ionic mobility also increases with temperature, thus increasing the conductivity of water and consequently the corrosion rate. An unusual temperature effect, known as thermogalvanic attack, is sometimes observed with copper alloys. Temperature differences of at least 65 °C between the ends of copper conduits will cause the cold end to be cathodic relative to the hot end. Consequently, copper will corrode at the hot end and its cations will migrate toward the cold end, where they will be reduced.

Increased water velocity can accelerate corrosion in various ways. Firstly, higher velocities accelerate corrosion by supplying dissolved oxygen to metal surfaces, or by establishing differential aeration cells that cause severe corrosion. On the other hand, if the flow rate is low, suspended substances may deposit on metal surfaces and cause under-deposit (crevice) corrosion. Severe flow rate might result in erosion and impingement attacks on metallic equipment. The combined outcome of these flow effects is that at sufficiently low chloride concentrations, the corrosion rate increases with water velocity to a maximum, and then declines (due to passivation) as the velocity is further increased.

There are various approaches to *mitigate* uniform corrosion: (i) selection of a more corrosion-resistant material, either metallic or nonmetallic; (ii) application of a coating as a barrier between the material and the environment; (iii) modification of the environment or addition of chemical inhibitors to reduce corrosion rate; (iv) application of cathodic protection.

17.6.2 Localized Corrosion

17.6.2.1 Crevice Corrosion

As the name indicates, crevice corrosion occurs in narrow spaces where an electrolyte can creep, usually by capillary forces, between two pieces of metal or between a metal and an insulator. It is a form of localized corrosion usually associated with a stagnant solution at the microenvironment level. As oxygen diffusion into the crevice is restricted, a differential aeration cell tends to be set up between the crevice (microenvironment) and the external surface (bulk environment). Oxygen depletion occurs in crevices, joints, and under corrosion deposits, metallic or nonmetallic deposits. Localized corrosion will occur at the area of low-oxygen concentration, which acts as an anode. Thus, crevice corrosion is often observed under riveted lap joints, bolt heads, gaskets, and O-rings. The severity of crevice corrosion susceptibility increases with increasing hysteresis of the polarization curve, i.e. the difference between E_b and E_p. Therefore, a higher

value of E_p reflects higher resistance to crevice corrosion. If the metal does not repassivate until a potential below the rest potential is reached, then it is very susceptible to crevice corrosion. Two concerns with this type of corrosion are that (i) the damage is concentrated within a small hidden area and thus inspection is very difficult, and (ii) the extent of damage cannot be monitored by mass loss/gain measurements.

Figure 17.21 illustrates the mechanism of crevice corrosion in a riveted joint in oxygen-containing seawater environment. The metal oxidation reaction is accompanied by oxygen reduction reaction. Initially, these reactions occur uniformly all over the surface, including in crevices; the oxygen concentration in the water occupying a crevice is equal to the level of soluble oxygen and is the same everywhere. However, after some time, oxygen in the crevice is depleted due to mass transport limitations; thus, its reduction in this area ceases. Since the area of the crevice is usually very small compared to the overall surface area, we will not observe any significant drop in the overall reduction rate; thus, the corrosion rate almost does not change. However, in order to compensate for the positive charge associated with metal cations being continuously formed within the crevice, anions diffuse into the crevice; these are often chloride ions. The chloride anion associates with the metal cation, and the metal chloride reacts with water to form metal hydroxide and hydrochloric acid. The pH in a crevice can reach very acidic values, sometimes equivalent to pure acids. The acidification of the microenvironment can lead to a significant increase in the corrosion rate of most metals. The dissociated chloride ion can react again with the metal ion, and the series of reactions repeat. This is an *autocatalytic process*.

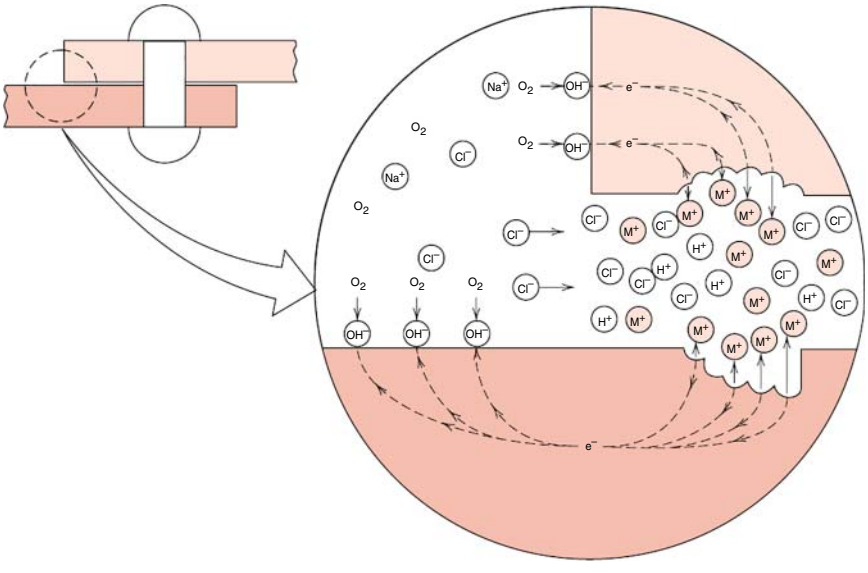


Figure 17.21 Schematic illustration of the mechanism of crevice corrosion between two riveted plates. Source: Reproduced with permission from M. G. Fontana, *Corrosion Engineering*, 3rd edition, McGraw-Hill Book Company, 1986.

The solid corrosion products seal the crevice microenvironment even further. As the anodic area is localized and small in comparison to the cathodic area, yet both should transfer the same current, the anode current density increases, together with the corrosion rate.

There are different ways to *control* crevice corrosion: (i) redesign equipment to eliminate crevices. (ii) Close crevices with non-absorbent materials or incorporate a barrier to prevent moisture penetration into the crevice. Close crevices in existing lap joints by continuous welding, caulking, or soldering. (iii) Design vessels for complete drainage; avoid sharp corners and stagnant areas. Prevent or remove builds-up of scale or other solids. (iv) Use one-piece or welded butt construction instead of bolting or riveting. (v) Select more corrosion-resistant or inert alloy. (vi) Stir the electrolyte. (vii) Reduce the conductivity of the electrolyte.

17.6.2.2 Filiform Corrosion

Filiform corrosion is a special type of crevice corrosion. In most instances, it occurs under protective (usually, organic) films; thus, it is often referred to as underfilm corrosion. The attack usually begins at imperfections in the coating, such as scratches and pores. The filaments are fine tunnels composed of corrosion products underneath the bulged and cracked coating. Filiform corrosion routinely occurs on coated steel cans, coated tin-plated steel, aluminum foil-laminated packaging, painted aluminum, painted magnesium, and other lacquered metallic items placed in areas subjected to high humidity. The attack occurs in most cases when the relative humidity is between 65% and 90%. The fluid in the leading head of a filiform is typically acidic, with a pH from 1 to 4. Oxygen (or air) and water are needed to sustain filiform corrosion, which indicates that a differential aeration cell is involved. CO_2 can stimulate the process by dissolving in water and forming carbonic acid. Chlorides, sulfate, and sulfide ions can dissolve in condensing atmospheric moisture and increase the attack. The optimum temperature for filiform attack is between 20 and 35 °C, with a corresponding relative humidity range of 60–95%. Above 95% humidity, blistering rather than filiform corrosion may occur (see ASTM D2803-09(2015)).

There are different approaches to *prevent* filiform corrosion: (i) reduce relative humidity to below 60% by the use of drying fans, dehumidifiers, or the addition of desiccants in packaging applications. (ii) Use more brittle films that crack and destroy differential aeration. (iii) Use impermeable coatings. Use zinc (galvanizing) and zinc primers on steels. Use zinc chromate primers, chromic acid anodizing, and chromate or chromate–phosphate conversion coatings on aluminum. (iv) Use multiple-coat paint systems. (v) Use less active metal substrates, such as copper, stainless steel, or titanium.

17.6.2.3 Pitting Corrosion

Pitting corrosion is a highly localized corrosion of a metal surface that is confined at a small area and takes the form of cavities. This is typically a process of local anodic dissolution, for example, at local breakdowns of the passive layer, where metal loss is exacerbated by the presence of a small anode and a large cathode. Pitting is commonly observed on surfaces with little or no general corrosion. It is often initiated at MnS inclusion sites, surface scratches, exposure of

dislocations or other defects to the surface, local breakdown of the passive layer due to exposure to chloride, bromide, and other halide ions, or local arbitrary variations in the composition of the electrolyte solution. An increase in the resistance to pitting corrosion is associated with an increase in E_b. While pitting will occur on a pit-free surface above E_b, it will occur only in the range of potentials between E_p and E_b if the surface is already pitted. Increasing the temperature usually increases the susceptibility to pitting. All metals that are susceptible to pitting corrosion are also susceptible to crevice corrosion, but not vice versa. Stainless steels are susceptible in particular. The pits usually grow in the direction of gravity. They can have different shapes, such as deep and narrow, shallow and wide, elliptical, undercutting, subsurface, and horizontal or vertical grain attack. ASTM G46-94(2013) describes the selection of procedures that can be used in the identification and examination of pits and in the evaluation of pitting, including a standard rating chart.

The mechanism of pitting corrosion is illustrated in Figure 17.22. In an alkaline chloride solution, anodic dissolution of stainless steel produces Fe^{2+}, which attracts negative anions such as Cl^- to the initiation site. Hydrolysis occurs following the reaction

$$Fe^{2+} + 2H_2O + 2Cl^- \rightarrow Fe(OH)_2 + 2HCl \tag{17.124}$$

Consequently, the pH drops in the initiation site. An autocatalytic mechanism of pit growth occurs. The acidic chloride solution accelerates anodic dissolution, which consequently concentrates more Cl^- in the pit. An insoluble cap of the corrosion product $Fe(OH)_3$ forms at the entrance to the pit, and Fe^{2+} diffuses

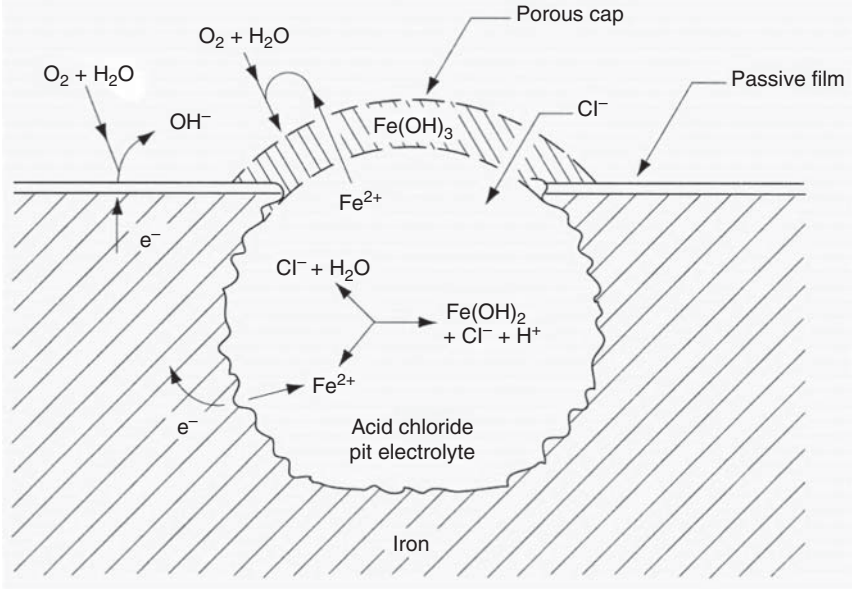


Figure 17.22 Schematic illustration of the processes occurring in an actively growing pit in iron. Source: Reproduced with permission from D. A. Jones, *Principles and Prevention of Corrosion*, Macmillan Publishing Company, 1992.

from inside to outside of the pit, where it is oxidized to Fe^{3+} and precipitates in the neutral bulk solution. The cap inhibits escape of Fe^{2+}, but is sufficiently porous to allow Cl^- migration into the pit, thus preserving high acidic chloride concentration within the pit. Measurement of the pH inside a pit is not an easy matter, but estimates based on various calculations and on measurements in model pits can lead to pH values as low as 1–2 for chromium-containing ferrous alloys, and about 3.5 for aluminum-based alloys, depending on experimental conditions. Anodic polarization of the inner pit occurs by coupling to the outer passive cathodic surfaces. Reduction of solute oxidizer such as oxygen consumes the electrons that are generated by the anodic reaction in the pit.

There are different ways to *mitigate* pitting corrosion: (i) increase the velocity of the fluid medium. (ii) Remove deposits of solids from exposed metal surfaces. (iii) Select alloys with higher alloying elements content; for example, in stainless steels the content of Cr should be at least 20 wt%, the content of Mo should be 2–6 wt%, and N may be added. Chromium blocks active surface sites, slows transport of aggressive anions, and lowers the flux of cation vacancies. Copper increases E_p of aluminum alloys as long as a second phase does not precipitate. (iv) Ensure high surface finishing.

17.6.3 Intergranular Corrosion

17.6.3.1 Sensitization

The term intergranular corrosion (IGC) refers to corrosion occurring at grain boundaries due to the presence of a second phase, contaminants, or atom segregation. A galvanic cell is established, with the anode around/at grain boundaries. Sensitization in austenitic stainless steels is one example of IGC. Austenitic stainless steels (which typically contain nominally 18 wt% Cr and 8 wt% Ni) are extensively used as construction material in chemical, petrochemical, fertilizer, and nuclear industries. One of the major problems associated with these steels is their susceptibility to IGC and intergranular stress corrosion cracking (IG-SCC) due to sensitization. When exposed to temperatures in the range 425–815 °C, $(Fe,Cr)_{23}C_6$ carbides precipitate at grain boundaries, causing chromium depletion around grain boundaries (see Figure 17.23). Consequently, these boundaries become susceptible to corrosion in chloride and caustic environments. Sensitization is often a consequence of improper heat treatment of fabricated components for stress-relieving purpose, prolonged service at elevated temperatures, and slow cooling from higher temperatures (e.g. hot working or during shutdown of plants operating at higher temperatures). ASTM A262-15 covers the standard practices for detecting susceptibility to IGC in austenitic stainless steels.

There are several ways to *mitigate* sensitization: (i) alloy the steel with Nb, Ta, or Ti, which have higher affinity to carbon than Cr and form carbides that are distributed more evenly in the microstructure. The concentration of Ti in 321 stainless steel, for example, is at least five times the concentration of carbon. The alloy 347 stainless steel contains both Nb and Ta; the concentration of the former should be at least 10 times the concentration of carbon. (ii) Reduce the carbon content to below 0.03 wt%, so that even if chromium carbides precipitate, their

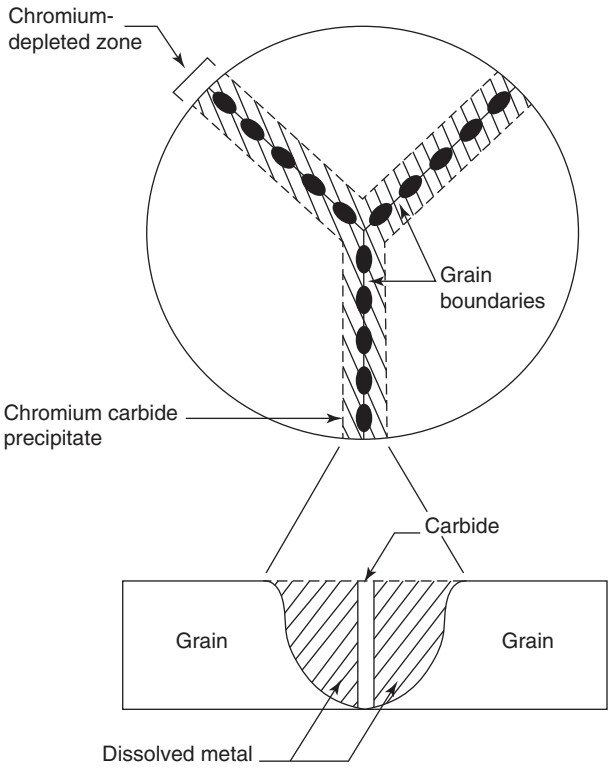


Figure 17.23 Schematic representation of carbide precipitation at grain boundaries and associated depletion of chromium in the vicinity of the boundaries, making the stainless steel susceptible to intergranular corrosion (sensitization). Source: Reproduced with permission from D. A. Jones, *Principles and Prevention of Corrosion*, Macmillan Publishing Company, 1992.

amount will not be sufficient to cause susceptibility to corrosion of grain boundaries. Such steels are designated with the letter "L," for example, 304L, 316L, and 317L stainless steels. (iii) Control the processing/joining parameters (e.g. in welding). (iv) Apply solution heat treatment at 1066–1121 °C (would dissolve existing carbides), followed by water quenching (fast cooling) so that there is not enough time for carbide reprecipitation.

17.6.3.2 Exfoliation

Exfoliation is a form of localized corrosion that primarily affects aluminum alloys in industrial or marine environments. Corrosion proceeds laterally from initiation sites on the surface and generally proceeds intergranularly along planes parallel to the surface. The corrosion products that form in the grain boundaries force metal away from the underlying base material, resulting in a layered or flake-like appearance.

Prevention is possible through proper alloy and temper selection. For example, commercially pure Al (1xxx series) and Al–Mn alloys (3xxx series) are quite resistant to exfoliation corrosion in all tempers. Tempers H116 and H117 have been suggested for Al–Mg alloys 5083, 5086, and 5456, rendering them highly resistant

to exfoliation corrosion. In Al–Zn–Mg alloys containing copper, such as 7075, resistance to exfoliation can be improved markedly by overaging (T7xxx temper for wrought products). In Al–Cu–Mg alloys (series 2xxx), artificial aging to the T6 or T8 conditions provides improved resistance.

17.6.4 Dealloying

Dealloying, also referred to as selective leaching or parting corrosion, is a corrosion process in which the more active metal is selectively removed from an alloy, leaving behind a porous weak deposit of the more noble metal. Specific categories of dealloying often carry the name of the dissolved element. For example, the preferential leaching of zinc from brass is called dezincification. In the case of gray iron, dealloying is called graphitic corrosion.

In the dealloying process, typically one of two mechanisms occurs: (i) alloy dissolution and replating of the cathodic element, or (ii) selective dissolution of an anodic alloy constituent. In either case, the metal is left spongy and porous and loses much of its strength, hardness, and ductility.

Copper–zinc alloys containing more than 15 wt% Zn are susceptible to dezincification. In the dezincification of brass, selective removal of Zn leaves a relatively porous and weak layer of Cu and copper oxide. Corrosion of a similar nature continues beneath the primary corrosion layer, resulting in gradual replacement of sound brass by weak, porous copper. Unless arrested, dealloying eventually penetrates the metal, weakening it structurally, and allowing liquids or gases to leak through the porous mass in the remaining structure. Dezincification is the usual form of corrosion for uninhibited brasses in prolonged contact with water high in oxygen and CO_2. Layer-type dezincification is easy to recognize visually. The original component shape and dimensions are usually preserved, but the metal changes from the gold yellow of zinc brass to the red of elemental copper at the dezincified surface. Plug-type dezincification is common on horizontal pipe/tube sections. Attack is usually confined to areas beneath deposits lying along pipe bottoms.

Prevention is possible through proper alloy selection. Brasses with copper contents of 85 wt% or more resist dezincification. Dezincification of brasses with two-phase structures is generally more severe, particularly if the second phase is continuous. Tin tends to inhibit dealloying, especially in cast alloys. Alternatively, the corrosion environment may be modified either by removal of harmful constituents or by addition of a proper inhibitor.

17.6.5 Galvanic (Bimetallic) Corrosion

Galvanic corrosion is an accelerated corrosion of a relatively active metal (anode) when it is brought in electrical contact with a more noble metal (cathode) in a common electrolyte. An example of this corrosion phenomenon is the increased rate of corrosion of steel in seawater when in contact with copper alloys, or the preferred dissolution of a zinc coating on a steel bolt, even when areas of the steel become exposed to the electrolyte. This type of corrosion is prevalent not

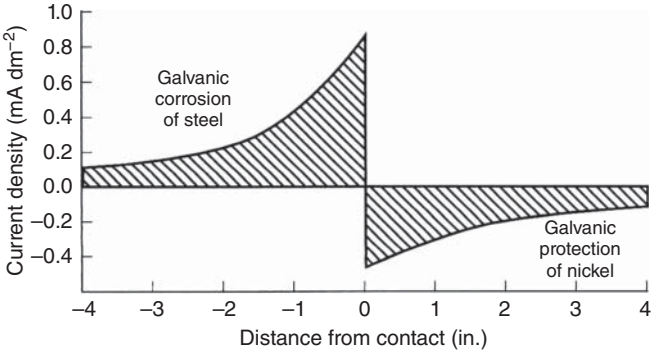


Figure 17.24 Effect of distance from the junction on the attack of the anode in a galvanic couple. Source: Reproduced with permission from H. R. Copson, *Trans. Electrochem. Soc*. **84** (1943) 71.

only in immersed structures or in pipelines carrying an ionically conducting liquid, but also in metals exposed to humid atmospheres, where the ionic path of conductivity is established through a thin film of moisture accumulating on the surface.

This form of corrosion may be either uniform or localized. It can be particularly severe under conditions where protective corrosion films do not form, or where they are removed by conditions of erosion corrosion. ASTM G71-81(2014) is a standard guide for conducting and evaluating galvanic corrosion tests in electrolytes, whereas ASTM G82–98(2014) is a common standard for construction and use of the galvanic series for predicting the galvanic corrosion performance. In Section 17.5.2.4 we described zero-resistance ammetry tests of galvanic couples. As explained in Section 17.3.2, the galvanic series is more appropriate than the EMF series (Table 17.1) for description of actual corrosion problems. Usually, the further apart two metals are in the series, the greater is the driving force for galvanic corrosion, and the greater the damage to the anode will often be. As a rule of thumb used by corrosion engineers, a potential difference of approximately 300 mV is required to initiate and sustain galvanic corrosion. With smaller potential differences, corrosion due to galvanic coupling is typically small and is commonly ignored. The anode-to-cathode surface areas ratio is also important: the lower this ratio is, the higher the current density on the anode is, and the higher the corrosion rate is. The maximal damage will occur near the contact between the two metals (see Figure 17.24), because the driving force farther along is diminished by the potential drop across the solution resistance. Low conductivity of the electrolyte solution and low polarization usually focus the galvanic attack closer to the contact zone. The nature of the environment is also important. In Figures 17.14 and 17.15 we have already used Evans diagrams to illustrate the effects of Fe–Zn coupling and the anode-to-cathode surface area ratio, respectively, on galvanic corrosion.

A number of procedures or practices can be used to *mitigate* galvanic corrosion: (i) select combinations of metals as close together as possible in the galvanic series suitable for the particular application or service environment. (ii) Avoid

couples in which the area of the more active material (anode) is relatively small. For example, a brass bolt in a steel structure is quite safe, whereas a steel bolt in a brass structure should be avoided. (iii) Completely insulate (electrically) dissimilar metals wherever practicable. If complete insulation cannot be achieved, consider using a material such as paint or plastic coating at the joint. (iv) Apply coatings with caution. Do not paint the less noble material without also coating the more noble material; otherwise, greatly accelerated attack may be concentrated at imperfections in coatings on the less noble metal. If only one surface can be painted, the more noble surface should be chosen to reduce or eliminate the cathode area. (v) If practical, add chemical inhibitors to the solution, to reduce its corrosiveness. (vi) Welded and brazed joints are better than threaded joints, as there is less risk of entrapment of moisture in grooves. (vii) Apply cathodic protection.

17.6.6 Environmentally Induced Cracking (EIC)/Environment-Assisted Cracking (EAC)

17.6.6.1 Hydrogen Embrittlement (HE)

The entry of atomic hydrogen into high-strength steels and most other structural alloys, either during fabrication (e.g. during machining, welding, pickling, or electroplating) or during service (e.g. due to corrosion processes, or cathodic protection), might result in loss of ductility, cracking, or catastrophic delayed brittle failures at applied stresses well below the yield strength of the steel. Sources of hydrogen atoms could be, for example, (i) hydrogen gas molecules that adsorb onto the metal surface and dissociate; (ii) recombination of hydronium cations, H_3O^+, with electrons at the metal surface; (iii) reaction between hydrogen-containing molecules such as alcohols and the metal (or oxide) surface to release hydrogen. Embrittlement can occur as a result of hydrogen located within the bulk of the alloy during the application of a load ("internal hydrogen embrittlement"). In addition, embrittlement might result from the exposure of an alloy under load to a hydrogen-containing environment ("external hydrogen embrittlement"). This phenomenon has received several names, including HE, and it often occurs in alloys that show no significant loss in ductility when measured by conventional tensile tests. Hence, several standard mechanical tests have been suggested to evaluate the susceptibility to either internal or environmental hydrogen (see Section 17.5.1). It is interesting to note that the amount of hydrogen that can cause severe embrittlement is minute, as low as several parts per million (ppm) by weight.

The HE process depends on three major factors: (i) the origin of the hydrogen; (ii) the transport processes involved in moving the hydrogen from its source to the locations where it reacts with the metal to cause embrittlement; and (iii) the embrittling mechanism itself. Several embrittling mechanisms have been proposed for crystalline materials. The most common HE mechanisms are the high-pressure bubble formation, reduction in surface energy (adsorption mechanism), reduction in the lattice cohesive force (decohesion mechanism), hydrogen interaction with dislocations, and hydride formation.

In high-strength steels (yield strength greater than 170 ksi), the effect of internal hydrogen supersaturation is significant. Internal HE by the mechanism of

high-pressure bubble formation might be extremely severe due to the ease of cavity (microcrack) growth. Degradation is most severe under static loads or at low strain rates, where high internal hydrogen pressures can be maintained within the cavity or microcrack. Since hydrogen interacts with the high-strength martensitic structure, fractographic evidence of hydrogen degradation by internal pressure formation often exists. Isolated areas are observed of either intergranular (IG) or transgranular (TG) fracture by brittle cleavage or interface separation, depending on the relative strength of the grain boundaries. These areas are separated by other areas exhibiting the normal fracture mode for rapid, overload failure in the alloy (e.g. microvoid coalescence or quasicleavage). In principal, the presence of hydrogen may also lead to the appearance of "hairlines" on the facets of some grains, or to a change in the size of microvoids. In Figure 17.25, the IG fracture is evident, together with hairlines and microvoids on the facets of the grains.

Hydrogen need not be present in concentrations above the solid-solution solubility limit to embrittle martensitic and ferritic steels. Both structures can be strongly degraded by relatively low concentrations of hydrogen originating either from the bulk of the alloy or from a hydrogen-containing environment. In general, the higher the strength level of the steel, the greater its susceptibility to HE. At low concentration of hydrogen, the exact mechanism of embrittlement is often difficult to establish. The fracture surfaces of smooth and notched tensile specimens often exhibit a brittle-to-ductile transition failure. At the region in contact with hydrogen, fracture is generally IG, occurring by interface separation of the grain boundaries. At the other regions of the specimen, failure represents the virgin material and may be of a ductile mode.

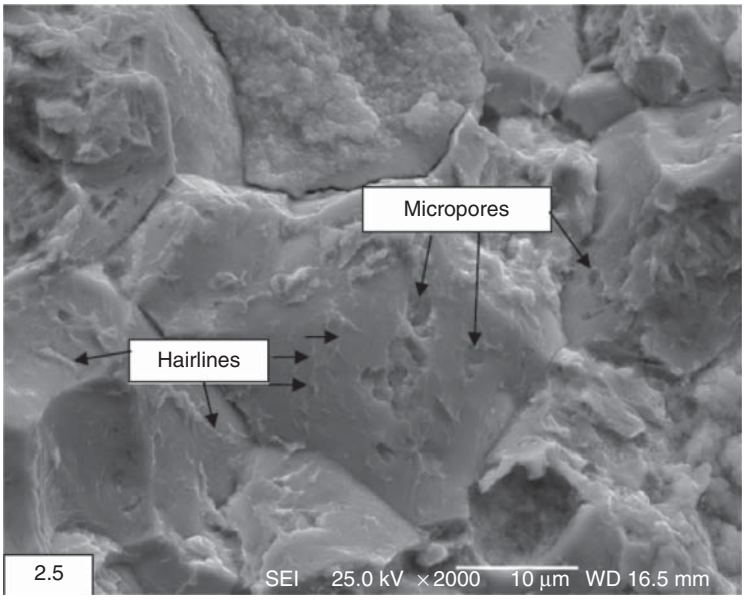


Figure 17.25 Fractography (scanning electron microscope image) of a steel part in aircraft landing gear that failed due to hydrogen embrittlement.

For a given hydrogen content, the tendency to embrittle increases with decreased strain rate. The embrittlement is more prevalent at room temperature; cracking decreases with increasing temperature and disappears almost entirely in steels above 200 °C (at high temperatures, hydrogen can be driven out of solid solution given a sufficient period of time).

Prevention of HE can be achieved by several means: (i) in the design stage, do not select materials with strength higher than what is actually required for proper functionality. (ii) During processing, maintain low hydrogen atmosphere (e.g. low-hydrogen welding rods stored in a dry place), and apply baking heat treatment whenever necessary before and/or after processes such as welding and electroplating. (iii) Natural hydrogen barriers, such as some oxides, seem to work well at low to moderate temperatures when the structure is lightly loaded (no surface deformation). At elevated temperatures, these coatings tend to degrade and become ineffective. (iv) Artificial hydrogen barriers, such as tungsten or glass, can be effective at reducing the rate of hydrogen permeation through a material. However, these coatings are difficult to apply, particularly on complex internal surfaces, and contain defects that allow for rapid hydrogen entrance at elevated temperatures. (v) Gas-phase inhibitors added to the hydrogen environment, such as oxygen, appear attractive in some applications, but such additions to fuel and coolant hydrogen are generally forbidden. In addition to HE, there are two other phenomena of hydrogen-induced damage: hydrogen-induced blistering and hydrogen attack.

17.6.6.2 Hydrogen-Induced Blistering

In the case of hydrogen-induced blistering, hydrogen is absorbed in the metal, diffuses inward, and precipitates as molecular hydrogen at laminations or inclusion/matrix interfaces. Consequently, enough pressure is built up to produce internal cracks. If these cracks are just below the surface, the hydrogen gas pressure in the cracks can lift up and bulge out the exterior layer of the metal so that it resembles a blister. The equilibrium pressure of the molecular hydrogen in the void, which is in contact with the atomic hydrogen in the surrounding metal, is great enough to rupture any metal or alloy. The way to *mitigate* this phenomenon is to minimize the amount of hydrogen at the surface, for example, by using corrosion inhibitors and avoiding cathodic protection and galvanic couples.

17.6.6.3 Hydrogen Attack

Hydrogen attack is an elevated-temperature phenomenon that affects carbon and alloy steels. Steel exposed to high-temperature, high-pressure hydrogen (e.g. in petrochemical plants) appears to be unaffected for days or months, but then suddenly loses its strength and ductility. It should be noted that hydrogen attack is different from HE. The former is irreversible damage, which occurs at elevated temperatures, whereas HE is often reversible and occurs at temperatures below 200 °C. In hydrogen attack of steel, the absorbed hydrogen reacts internally with carbides to form methane bubbles along grain boundaries; these bubbles subsequently grow and merge to create fissures, decarburization is evident along grain

boundaries, and unexpected failure eventually happens. Steel composition, exposure time, and stress level will determine the extent of attack. The only practical way to prevent hydrogen attack is to use steels that have been found to be resistant to this type of deterioration in the specific environment. Among others, the carbon content in the steel should be minimized.

17.6.6.4 Stress Corrosion Cracking (SCC)

SCC is a phenomenon associated with the combination of static tensile stress and corrosive environment. Failures often occur in mild environments under tensile stresses well below the macroscopic yield strength of the material. The origin of tensile stresses may be external forces, thermal stresses, or residual stresses. In general, as these stresses increase, the time required to initiate SCC decreases. The kinetics of SCC also depends on the chemical and metallurgical state of the material (e.g. chemical composition, thermal condition, grain size, presence of secondary phases and precipitates); on environmental conditions (e.g. environment composition, temperature, pressure, pH, electrochemical potential, solution viscosity, and mixing); and on crack geometry and stress state (uniaxial, triaxial, etc.).

SCC may be either IG or TG (or mixed), depending on the alloy, its microstructure, and the environment. However, the crack follows a general macroscopic path that is always normal to the tensile component of stress. TG failures are less common than the IG ones, but both may exist in the same system, or even in the same failed part, depending on conditions. The IG failure mode suggests some inhomogeneity at the grain boundaries. For example, segregation of sulfur and phosphorus at grain boundaries has been suggested as the cause of IG SCC of low-alloy steels.

Despite a large amount of data, the mechanisms of SCC are still under considerable debate. Several mechanistic models have been suggested for crack *propagation*. These mechanisms are divided into anodic and cathodic SCC.

Anodic SCC: For dissolution-based mechanisms it is assumed that localized anodic dissolution of the metal takes place at the crack tip, when it becomes possible according to the Pourbaix diagram (or a modified form of it, to adjust for the current conditions), and that the stress simply helps the corrosive environment to reach the metal. The crack velocity is thus related to the rate of metal dissolution. The main mechanisms are the active path mechanism, the film rupture model, the slip dissolution model, the corrosion tunnel model, and the coupled environment fracture model. All of these models assume that crack propagation is related to the anodic dissolution of the metal after rupture of a passive film by localized strain. However, suggestions have been made that these strains are in fact not enough to break the film. They also fail to explain TG SCC, which, unlike IG SCC, does not follow a chemically active path. Recent theories have focused on the discontinuous nature of crack propagation and the cleavage features observed on the fracture surfaces, leading to the development of cleavage models. These are the tarnish rupture mechanism, the film-induced cleavage mechanism, the adsorption-induced cleavage mechanism, and the surface mobility mechanism.

The slip dissolution model was proposed as a refinement of the anodic dissolution model. In the case of sufficient disposal of oxygen, a passive film is assumed

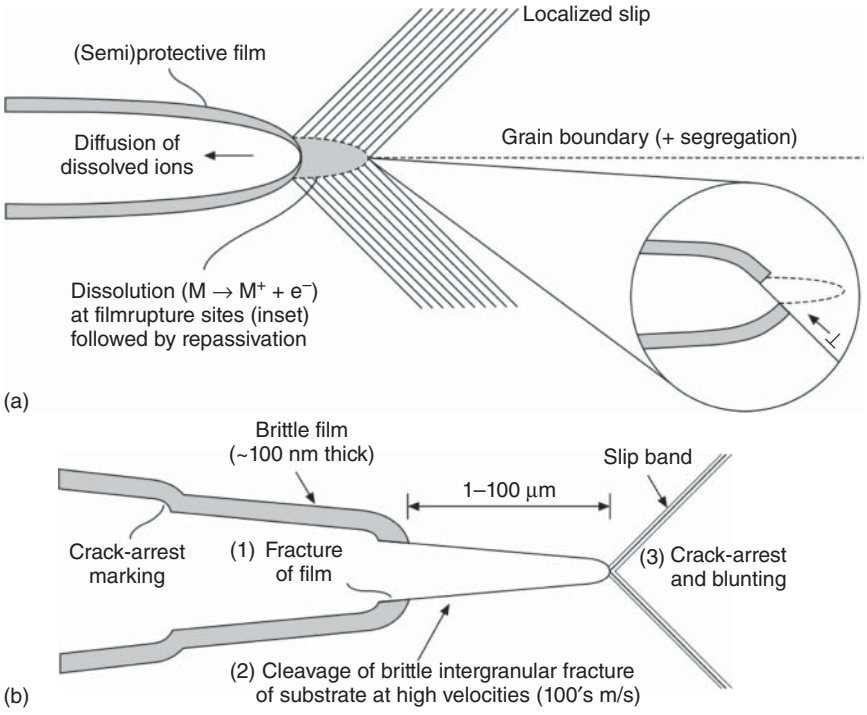


Figure 17.26 Schematics of: (a) slip dissolution involving rupture of the oxide film at the crack tip due to localized slip, followed by repassivation behind crack tip, and (b) film-induced cleavage. Source: Reproduced with permission from S. Lynch, *Corros. Rev*. **30** (2012) 63.

to form over the crack walls and the crack tip. It stops the crack growth until the buildup of plastic deformation or creep ruptures the passivating film by slip bands that intersect the crack, exposing bare metal at the crack tip. This metal dissolves intermittently in the corrosive medium, giving rise to crack propagation, until repassivation occurs (see Figure 17.26a). Alternatively, the crack tip may remain film-free if the strain rate is high compared with the repassivation rate. A related proposal is the film-induced cleavage model (Figure 17.26b) that assumes the formation of an environmentally induced brittle film at the crack tip. When it fractures, the crack continues into the substrate for distances greater than the film thickness. One complication arising from this model is that the observed crack geometry and the relatively little corrosion observed on the fracture surface are inconsistent with the depth of corrosion and plastic deformation at each film rupture cycle. Austenitic stainless steels in chloride-containing environments are often considered to undergo SCC via an anodic dissolution mechanism.

Cathodic SCC: These hydrogen-related mechanisms are based on the observations that a few parts per million of hydrogen, formed by electrolysis of water, might cause HE in high-strength steels and alloys. There has been some evidence, for example, that the solution within the micro-volume of the crack becomes acidified, probably by hydrolysis reactions similar to those that occur in pits.

Figure 17.27 The river branching pattern typical of SCC (https://corrosion.ksc.nasa.gov/stresscor.htm).

Therefore, the presence of hydrogen in the crack and the brittle cleavage characteristics of TG cracks have prompted support for the HE mechanism, at least for certain high-strength materials. The transport of hydrogen to the degradation sites, accumulation and trapping at microstructural heterogeneities, e.g. dislocations and grain boundaries, is an important step in the HE process. Mechanisms of HE related to hydrogen-induced atomic decohesion, hydrogen-induced phase transformation, hydrogen-enhanced localized plasticity (HELP), and hydrogen-enhanced stress-induced vacancies (HESIVs) have been suggested. Martensitic steels are usually more prone to cathodic SCC.

Cross-sections of SCC frequently show branched cracks (see Figure 17.27). This river branching pattern is unique to SCC and is used in failure analysis to identify when this form of corrosion has occurred.

The resistance to SCC depends, among others, on the thermal condition of the alloys as well as on the crystallographic direction in the material. For example, on a scale of 1–4, where 1 denotes the highest resistance to SCC, the resistance of wrought alloy 2024 is 3 in the T4 and T3 temper conditions, and 2 in the T6 and T861 conditions. The resistance of a plate of wrought Al 2024 in the T4 condition is 1 in the longitudinal (L) test direction, 2 in the long transverse (LT) test direction, and 4 in the short transverse (ST) direction. In contrast, the resistance of Al 7075-T73 is 1, regardless of the test direction. On the anodic polarization curve, two zones of the highest likelihood of TG-SCC are around the bottom and upper parts of the passive region. IG-SCC may occur within a broader potential range, due to inhomogeneities at the grain boundaries.

There are various ways to *prevent* SCC:

1) *Mechanical*:
   1.1 Avoid stress concentrations, such as sharp corners.
   1.2 Release residual stresses, for example, by applying annealing or stress relief heat treatments. Assure proper solidification rates during casting or welding to prevent residual stresses.

1.3 Introduce compression stresses to the surface, e.g. by shot peening.
1.4 Ensure good surface finish.
1.5 Lower service stresses.
1.6 Use non-destructive tests (NDT) when cracks are unavoidable.

2) *Metallurgical*:
2.1 Use a metal that is resistant to SCC in the specific environment. For example, carbon steels are better than stainless steels; using low carbon steels is preferable.
2.2 Change the microstructure of the alloy. For example, large grains will make the alloy more prone because they are associated with lower yield strength. Minimize prone interfaces such as grain boundaries, dendrites, and inclusions. Quenching and tempering, as well as HAZ might cause higher crack growth rates. Thermal treatment may redistribute atoms at the grain boundaries, thus reducing IG-SCC.
2.3 Use metallic or conversion coatings. Conversion coatings on Al, Mg, and Ti are more effective when combined with high-quality paints.

3) *Environment*:
3.1 Control the environmental conditions, when possible.
3.2 Apply cathodic or anodic protection. In cathodic protection, the potential must be lower than the minimal causing dissolution, but higher than the maximal causing HE. Anodic protection is not useful when the cracking is associated with pitting, because this type of protection requires an environment that enables stable passivation layers at potentials above the cracking domain.
3.3 Add inhibitors (e.g. phosphates). Some inhibitors will increase the potential above the SCC zone, while others will prevent hydrogen absorption in the metal. Some inhibitors are impregnated in paints or organic coatings, preferably in the priming layer.
3.4 For alloys with passivity regime, allow oxygen in the environment to renew the oxide layer.
3.5 Apply organic coating.
3.6 Change the temperature, if possible.

17.6.6.5 Corrosion Fatigue (CF)

The term CF refers to the initiation and propagation of cracks as a result of the synergistic effect of a cyclic stress and a corrosive environment. An aggressive environment usually has a harmful effect on fatigue strength and fatigue lifetime, producing failure in fewer stress cycles than would be required in a more inert environment. The action of the cyclic stress leads to local destruction of surface oxides, and hence, corrosion pits form. The cyclic stress also removes corrosion products that usually inhibit further corrosion. The bottom of the pit is anodic compared to the remaining metal; therefore, corrosion propagates inward, being helped by destruction of oxide layers by cyclic strain. Cracking will appear when the pit becomes sufficiently sharp to establish high stress concentration.

CF depends strongly on the interactions between mechanical, metallurgical, and environmental variables. Mechanical variables include the maximum stress or stress intensity factor, stress amplitude, stress ratio, cyclic loading frequency,

state of stress, residual stress, crack size, and shape and their relation to component size and geometry, etc. Metallurgical variables include alloy composition, distribution of alloying elements and impurities, microstructure and crystal structure, heat treatment, mechanical working, preferred orientation of grains and grain boundaries (texture), mechanical properties (e.g. strength or fracture toughness), etc. Environmental variables include temperature, pH, types of environment (e.g. gaseous or liquid), partial pressure of damaging species in gaseous environments, concentration of damaging species in aqueous or other liquid environments, electrical potential, viscosity of the environment, coatings, inhibitors, etc.

CF usually leads to TG fracture with little or no branching. Individual cracks may occur, although families or parallel cracks are more common. Decreasing the strain rate increases the effect of environment. CF is more pronounced at low stress frequencies. Often, endurance limit is not evident in aggressive environment.

All metals and alloys are susceptible to CF. Even some alloys that are immune to SCC, for example, ferritic stainless steels, are subject to failure by CF. *Remedies* include reduction or elimination of cyclic stresses, selection of a material or heat treatment with higher CF strengths, reduction or elimination of corrosion, or a combination of these procedures. Shot peening is usually effective in prolonging fatigue life in air by introducing residual compressive stresses in a metal surface, thus reducing the mean tensile stress at potential crack-initiation sites. In more aggressive environments, however, shot peening may have only limited value, because general corrosion can eventually remove the surface layer and, thus, the beneficial compressive residual stresses. Nitriding, which also introduces compressive residual surface stresses, can improve the CF resistance of steels, particularly when a relatively short life is required. Cathodic protection is helpful in reducing the influence of corrosion on fatigue of metals and alloys exposed to aqueous environments, except in such alloys as high-strength steels that are subject to hydrogen-induced delayed cracking. Anodic protection has been used in the case of stainless steels. Inhibitors are sometimes added to the environment or included in organic coatings to eliminate CF. In most instances, the environment cannot be prevented from contacting the component; therefore, coatings are helpful.

17.6.7 Erosion Corrosion

In erosion corrosion, the corrosion rate is accelerated due to the relative motion of a corrosive fluid and a metal surface. It is particularly severe in copper alloys. The increased turbulence caused by pitting on the internal surfaces of a tube can result in rapidly increasing erosion rates, eventually resulting in a leak. Erosion corrosion can also be aggravated by faulty workmanship. For example, burrs left at cut tube ends can upset smooth water flow, and cause localized turbulence and high flow velocities.

Remedial practices include the following: (i) make adequate material selection (note that high hardness does not necessarily guarantee high resistance). (ii) Reduce fluid velocity. (iii) Promote laminar flow (e.g. increase pipe diameters).

(iv) Avoid rough surfaces. (v) Avoid abrupt changes in flow direction. (vi) Direct tank inlet pipes away from the tank walls. (vii) Align carefully welded and flanged pipe sections. (viii) Impingement plates of baffles designed to bear the brunt of the damage should be easily replaceable. (ix) Use filtration or settling to remove abrasive particles. (x) Use water traps in steam and compressed air systems to decrease the risk of impingement by droplets. (xi) Use deaeration and corrosion inhibitors. (xii) Apply cathodic protection. (xiii) Apply protective coatings.

17.6.8 Microbiological Corrosion (MIC)

The term MIC refers to corrosion in the presence of microbial species, including bacteria and the associated biofilms produced by those bacteria. MIC occurs on a variety of metals and alloys such as iron and steel, copper, nickel, aluminum, titanium, and their alloys, as well as on nonmetallic materials such as concrete and polymers. MIC is a common problem in industrial processes due to the presence of microbes, adequate nutrients, and corrosive byproducts. MIC failures have been reported for mild steel piping and equipment exposed in the marine environment, oil refining industry, nuclear power plants, etc.

The term "biofouling" is often used to describe the undesired development of microbial layers (i.e. biofilms) on surfaces. Biofilm development begins when "planktonic" bacteria, i.e. unattached individual cells, adhere to a surface. As adherent cells grow and divide, proximity to the surface induces physiological adaptations, including secretion of exopolysaccharides (EPSs) to create a protective matrix surrounding the cells. This EPS matrix acts as a barrier in which diffusive transport prevails over convective transport. Hydrated EPS contributes to the bulk of the volume of a biofilm, and is primarily responsible for its slimy macroscopic properties. The thickness of a biofilm can vary greatly according to the environmental conditions. In systems with high shear forces, biofilms may grow only up to a few micrometers within the laminar boundary layer, while in microbial mats or in sewage systems, biofilms can reach the thickness of many centimeters. In most natural flowing systems, the average biofilm thickness ranges from 50 to 500 μm.

Bacteria present in the aqueous medium often have the potential to increase or decrease oxygen transport to the surface; consequently, these organisms have a role in increasing or decreasing general corrosion. Most MIC, however, manifests itself as localized corrosion because most organisms do not form in a continuous film on the metal surface. Biological organisms fall under two groups based on the type of corrosion they cause – either anaerobic or aerobic corrosion. Sulfate-reducing bacteria (SRB) from the genera *desulfovibrio* and sulfur-oxidizing bacteria of the type *thiobacillus* are examples of anaerobic and aerobic bacteria, respectively. Colonies of aerobic bacteria (i.e. that consume oxygen for their development) that are established on the surface form concentration cells, similar to those encountered in crevice corrosion. The anaerobic SRB is a group of bacteria that includes more than 220 species that produce H_2S and use sulfates as the terminal electron acceptor. Anaerobic means that the cells cannot metabolize and/or replicate in the presence of oxygen, although many species can temporarily tolerate low levels of oxygen. Furthermore,

anaerobic conditions capable of supporting SRB growth can be created in overall aerobic environments, due to the microniches created within the bacterial biofilm/corrosion product layer.

Figure 17.28 illustrates the mechanisms by which a biofilm can contribute to corrosion: (A) creation of anaerobic zones, (B) concentration of corrosive chemicals, (C) concentration of ferrous ions, (D) electron conduction away from surface, (E) creation of differential aeration zones, (F) binding of corrosion promoters, and (G) disruption of passivating film.

The Titanic shipwreck, which has been submerged 4 km below the surface of the North Atlantic Ocean for nearly 100 years, is a famous example of MIC (and corrosion at great depths, in general). As the pH of seawater near its surface is around 8.1, oxygen reduction (and not hydrogen reduction) should be the cathodic reaction in corrosion processes. Since the concentration of oxygen in great depths is very low, one could predict that metal parts of submerged shipwreck would corrode very slowly. However, as the Titanic taught us in 1985, this is not the case!

Below 200 m, there is not enough light for plant growth. Dead organisms provide nutrients for deep-dwelling animals and microorganisms, while thermal vents from the ocean floor release sulfur-rich water. Ocean temperature decreases with depth as pressure increases. Gas solubility increases while salt solubility decreases as temperature decreases. At 4 km below the ocean's surface, with the water temperature being near freezing and little oxygen (~0.2 ppb) or salt dissolved in the water, any chemical reaction would be hindered. However, at great depths of the ocean (below 1 km), live the anaerobic SRB, also known as "iron-eating" bacteria. Equation (17.125) shows the reduction reaction caused by the SRB, which reduces the oxidation number of sulfur from +6 to −2:

$$SO_4^{2-} + 10H^+ + 8e^- \rightarrow H_2S + 4H_2O \tag{17.125}$$

Going deeper in seawater, the solubility of CO_2 increases, causing the water to become more acidic, as shown in Eq. (17.126):

$$H_2O + CO_2 \rightarrow H_2CO_3 \tag{17.126}$$

Water in some areas surrounding the Titanic was found to be of a pH as low as 4. Consequently, the corrosion of iron occurred more rapidly in this acidic water. Also, the higher concentration of hydrogen cations, made available by the dissolved hydrogen sulfide in the water, contributes to the corrosion of iron. The SRB are able to convert H_2 into H^+, which then chemically reacts to reduce the sulfate ions in the water, and more iron is consequently ionized. The relevant chemical equations are

$$H_2S \rightarrow 2H^+ + S^{2-} \tag{17.127}$$

$$Fe(s) + 2H^+ \rightarrow Fe^{2+} + H_2(g) \tag{17.128}$$

$$Fe^{2+} + S^{2-} \rightarrow FeS(s) \tag{17.129}$$

H_2S is a weak acid that can react with almost any metal, except gold and silver. Iron sulfide forms insoluble deposits similar to rusticles, although they have

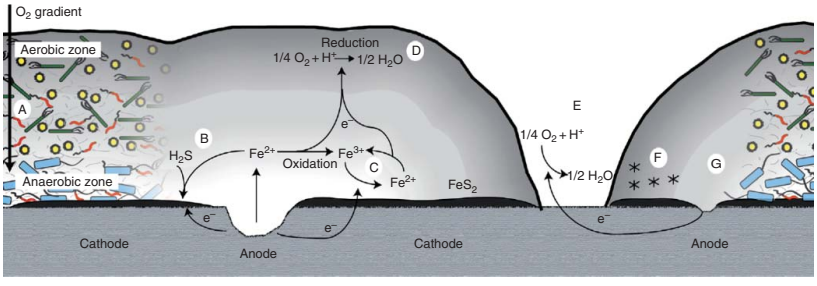


Figure 17.28 Mechanisms by which a biofilm can contribute to corrosion. Source: Reproduced with permission from K. Li, M. Whitfield, and K.J. Van Vliet, *Corros. Rev*. **31** (2013) 73.

a black and less porous appearance than rust. The precipitation of iron sulfide isolates more H^+ ions, making them available for involvement in the oxidation of more iron. This, in turn, increases the rate at which the steel hull of the Titanic rusts. Another factor in the increased corrosion rate of the Titanic is the organic material aboard the ship, namely, the interior wooden paneling. As the cellulose in the wood broke down, the released oxygen further contributed to the corrosion of the iron hull. This released oxygen not only provided a direct route for the cathodic reaction, but it also fed aerobic bacteria, which produced byproducts that decreased the pH of the water surrounding the Titanic.

Several procedures have been studied to *mitigate* MIC: (i) in the design stage, avoid stagnant zones. (ii) Use corrosion resistant metals. (iii) Apply protective coatings. (iv) Keep the system clean as much as possible. (v) Use pigging (i.e. mechanical delamination of corrosion products). (vi) Apply anodic or cathodic protection. (vii) Add biocides to the solution. Biocides are chemicals that kill or control microorganisms. Most of them are cationic materials that act at the surface. (viii) Introduce competing bacteria. (ix) Add or remove nutrient (e.g. via addition of nitrates, or removal of sulfates).

17.7 Corrosion Protection

There is a variety of corrosion control strategies, including the following: (i) selection of a more corrosion-resistant (more inert) alloy. Replace metallic with nonmetallic materials, whenever possible. (ii) Heat treatment (e.g. stress relief). (iii) Selection of proper manufacturing processes (e.g. welding instead of riveting). (iv) Change of the product design (geometry, matching surfaces, avoid crevices and stagnant zones, prevent turbulence, etc.). (v) Utilization of coatings (organic, inorganic, and metallic) and paints. (vi) Change of environment. For example, add corrosion inhibitors and biocides, and remove salts and aggressive components such as oxygen (deaeration), hydrogen sulfide, chlorides, and ammonia. Dry the atmosphere (removal of water vapors). (vii) Change in operating conditions (pH, lower temperature, etc.). (viii) Electrochemical – applying cathodic or anodic protection. (ix) Use of proper maintenance procedures (oil and greases, drying, etc.).

17.7.1 Cathodic Protection

CP is the most widely applied electrochemical corrosion control technique. This is accomplished by applying a direct current to the structure, which causes the potential to change from the corrosion potential (E_{corr}) to a protective potential in the immunity region. The required cathodic protection current is supplied either by sacrificial anode materials or by an impressed current system.

17.7.1.1 Cathodic Protection with Sacrificial Anodes

Cathodic protection with sacrificial anodes can be viewed as a form of galvanic corrosion, put to good use. In this case, an active metal (most often zinc, but also

Figure 17.29 Ductile iron municipal water main protected with tape wrap and sacrificial magnesium anodes. Source: Reproduced with permission from P. R. Roberge, *Handbook of Corrosion Engineering*, 2nd edition, McGraw-Hill, 2012.

magnesium or aluminum) is employed as a sacrificial anode. It is attached to the steel structure being protected in one or several locations, but does not constitute part of the structure itself (see Figure 17.29). The steel structure becomes the site of the cathodic reaction, and its potential is driven in the negative direction so that the rate of corrosion is slowed down to an acceptable level. The sacrificial anode, of course, corrodes at a relatively high rate and must be replaced periodically, but damage to the structure by corrosion can be minimized.

In addition, chemical backfill is widely used to surround galvanic anodes when they are buried in the earth. There are several advantages to this. The special backfill provides a uniform environment to promote uniform anode consumption and maximum efficiency. It isolates the anode material from direct contact with the earth, thus preventing negative effects by soil minerals that might otherwise build up high-resistance films on the anode surface. It has a low electrical resistivity that provides a lower anode-to-earth resistance and greater current output for cases where the surrounding soil is of higher resistivity.

Sacrificial anodes are most commonly employed to protect the hulls of ships and boats, for offshore oil rigs, and underground pipelines. They are designed to be replaced, if necessary, during routine maintenance. But, these are not their only uses. Galvanized steel, which is coated with a thin layer of zinc, functions in a similar manner. The coating provides a certain degree of protection as long as it is intact. When it is partially removed, by either corrosion or abrasion, the exposed surface is still protected cathodically and does not corrode. In contrast, a Ni/Cr coating may provide much better protection as long as it is intact, but there will be severe local galvanic corrosion in the area of a scratch, since the coating will act as a large-area cathode and the exposed steel surface as a small-area anode – a very bad combination.

The theory of cathodic protection is simple and straightforward. The real engineering challenge is to design the anodes and position them so that they provide uniform current distribution on the part being protected. There are two aspects to this problem: parts of the structure that are too far away from the anode, or screened from it, may not have sufficient protection, and parts that are too close to the anode may be "overprotected." We recall that the cathodic reaction in most cases is hydrogen evolution or oxygen reduction. Both reactions lead to the formation of OH^- ions near the surface. This can weaken the bonding of paints and other nonmetallic coatings to the surface, causing delamination. Also, excessive hydrogen evolution can lead to penetration of atomic hydrogen into steel, causing embrittlement, which could lead to devastating structural damage. Thus, under certain circumstances, overprotection can be worse than no protection at all.

The advantages of cathodic protection with sacrificial anodes include the following: (i) most metals in contact with an aqueous environment having a near-neutral pH can be cathodically protected. (ii) No external power sources are required. (iii) Installation is easy (and requires low costs). (iv) Cathodic interference is unlikely in other structures. (v) System requires low maintenance. (vi) System is essentially self-regulating. Less inspection and archiving of data is required compared to impressed-current cathodic protection (ICCP).

Unfortunately, these relatively simple systems also have some limitations: (i) current and power output are limited. (ii) High-resistivity environments or large structures may require excessive number of electrodes. (iii) Anodes may have to be replaced frequently under high current demand. (iv) Anodes can increase structural weight if directly attached to a structure. (v) A protected structure should be electrically isolated from its surrounding.

There are no simple equations from which the current distribution on structures having complex geometries may be obtained. A number of numerical methods have been developed, and a wealth of practical experience is available to overcome this problem. The real problem lies in the fact that conditions during service life change, not always in a predictable manner. Thus, the protective coating on the hull of a ship or on a pipeline may be damaged with time, changing the effective area that needs to be protected. Rain or drought will change the conductivity of the soil in which a pipeline is buried. The supply of oxygen to the immersed parts of a boat changes dramatically when it lifts anchor to sail at full speed. The conductivity of the water changes by several orders of magnitude when a ship sails from a river into the open sea. All these factors change the current distribution. If designed for optimum protection at sea, a ship will be overprotected in fresh water, and vice versa.

The best way to overcome the limitations of cathodic protection related to changes in the environment is to monitor the potential and adjust the cathodic currents accordingly. This cannot be readily done with sacrificial anodes, and the method of ICCP is sometimes preferred, in spite of its higher cost.

17.7.1.2 Impressed-Current Cathodic Protection (ICCP)

ICCP entails the use of an external power source in combination with a stable anode (graphite, Pt-coated Nb, titanium for protecting steel in concrete, Pb–Ag in seawater environment, etc.). The potential of the specimen being

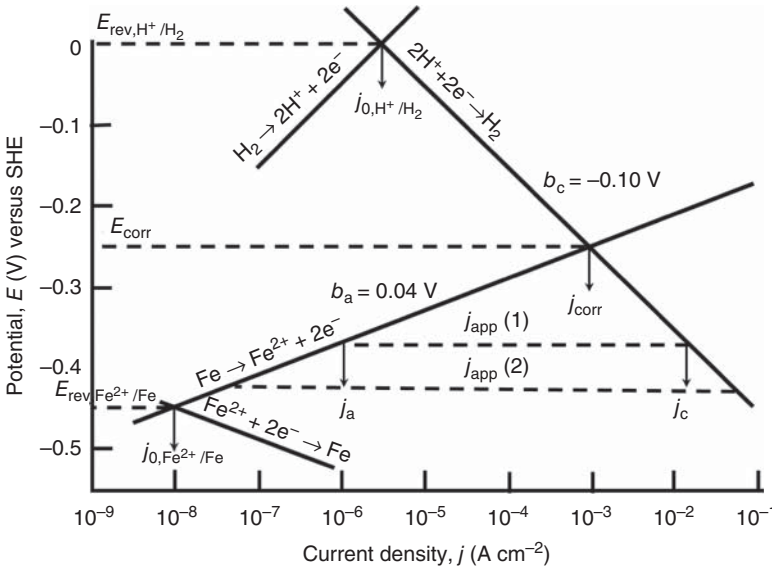


Figure 17.30 Evans diagram for iron corrosion in an acidic deaerated environment, showing the principle of ICCP. Two dashed horizontal lines show two levels of cathodic protection. The impressed current is the difference between the cathodic and the anodic currents at this potential.

protected is forced to negative values, and its rate of corrosion is consequently reduced. The result of impressing a cathodic current on the structure is shown in Figure 17.30. For the parameters used to draw this figure we obtain $E_{corr} = -0.25$ V versus SHE and $j_{corr} = 10^{-3}$ A cm^{-2}. Applying a cathodic current density $j_{app}(1) = j_c - j_a \cong 10^{-2}$ A cm^{-2} shifts the potential to -0.37 V, and the anodic current density (which represents the rate of corrosion of the cathodically protected surface) is reduced to $j_a \cong 10^{-6}$ A cm^{-2}, as shown by the dashed line (1). The current flowing in the external circuit is the difference between the cathodic and the anodic currents flowing at this potential. Thus, application of a current of c. 10^{-2} A cm^{-2} causes a decrease of the corrosion rate by 3 orders of magnitude. The rate of corrosion can be reduced further by increasing the applied cathodic current, as shown by line (2) in Figure 17.30.

It should be noted that using the ICCP does not alleviate the need to calculate the current distribution and locate the anodes in a way that ensures the best uniformity of current distribution on the protected structure. Furthermore, corroding structures do not have uniform corrosion potentials or protection requirements over their entire surface. Practical protection criteria need to take such variations into account. A protection criterion that is widely used for buried steel structures under aerobic conditions is to maintain a structure potential of 850 mV versus saturated $Cu/CuSO_4$ reference electrode.

ICCP is usually not used by itself. ICCP of ships is always used in conjunction with protective coatings. The coatings are intended as the primary protection, and the ICCP is a backup in those areas where coating defects may be present. The potential can be monitored continuously at different locations, and the impressed

current can be adjusted automatically in order to maintain the desired level of corrosion protection. In the period immediately subsequent to the application of the coating, there is very little demand on the impressed current system. As the coating deteriorates during the operational life of the ship, the CP current demand increases. Eventually, the current demand placed may exceed the capabilities of the design, with high anode currents causing even more damage to the coating, particularly in areas adjacent to the anodes, where the concentration of hydroxyl ions may be very high. In addition, the high local currents around the anodes may reduce the protection supplied to the rest of the structure. On a supertanker, the initial current of 10 A may rise to over 1000 A during the course of its operational life.

ICCP requires a little more sophistication than the use of sacrificial anodes, but it also lends itself to periodic adjustment and provides higher flexibility, particularly when structures having rather intricate shapes are considered. Other advantages of ICCP are that it can be used for all metals, corrosion rate can be reduced to zero, the current output is controllable, high currents can be applied when necessary, and non-insulated structures can be protected. However, there are also some other limitations: voltage is continuously applied, installation requires more qualified personnel, or electrical disorders (such as alternating current and stray current) might disturb the normal working of a cathodic protection system. Stray currents exist in the soil due to electric tractions, arc welding, galvanizing plants, natural electric currents (telluric currents) in the Earth's crust, etc. The intensity of such stray currents may reach several hundred amperes. If a metal is dug into the soil, a stray current can enter into this structure and then leave it somewhere. At the sites where electric current enters or leaves a metal, chemical reactions must take place. Therefore, the stray current forms anodic and cathodic sites on the surface of the same buried structure. Anodic sites would force the protection potential in the positive direction, which might shift the potential out of the thermodynamic immunity/nobility (see Figure 17.3).

17.7.2 Anodic Protection

Anodic protection is used to protect metals and alloys with active–passive behavior in the specific service environment. Here, the metal potential is shifted to the passive region from the active region by the application of a direct current. In cases where it is applicable, anodic protection has great advantages over cathodic protection, for a number of reasons. First, it requires typically only 1–2 $\mu A\ cm^{-2}$ (often, a direct measure of the corrosion rate during protection), about 2 orders of magnitude less than cathodic protection. In addition to the saving in energy, the negative side effects of cathodic protection, namely, HE and delamination of nonmetallic coatings resulting from the high pH generated near the cathodic sites, are eliminated. Since anodic passivation is performed potentiostatically, and since the currents involved are very small, uniform current distribution is easier to maintain, and overprotection is not likely to occur.

Unfortunately, anodic protection is limited to certain environments where the liquid in contact with the protected structure is well defined and known to allow passivation. There is, however, a widely used chemical form of anodic protection,

entailing paints that contain strong oxidizing agents (anodic inhibitors). The corrosion protection afforded by such paints can be the result of a number of mechanisms operating in parallel. One of them is the high positive potential set up by the oxidizing agent, bringing the metal into the passive region.

Anodic protection is adequate for moderate-to-aggressive solution corrosiveness, including strong alkaline or acidic environments. The throwing power is very high (compared to cathodic protection where it is low), and the operation cost is very low. On the other hand, installation and maintenance costs are high.

17.7.3 Corrosion Inhibitors

The use of corrosion inhibitors is one of the most common practices of corrosion prevention. One unique advantage of inhibitors is that their addition can be implemented without disruption of a process. There are hundreds of different inhibitors in commercial use. Corrosion inhibitor, as defined by ISO 8044:2015, is a "chemical substance that, when present in the corrosion system at a suitable concentration, decreases the corrosion rate, without significantly changing the concentration of any corrosive agent." Typical concentrations of inhibitors are 10–50 mg L^{-1} in open systems and thousands milligrams per liter in closed systems. Corrosion inhibitors can cause changes in the state of the protected metal surface through adsorption or formation of compounds with metallic cations. This results in a reduction of the exposed surface area of the metal and an increase in the activation energy of the corrosion processes. The adsorption and formation of protective layers on metals is greatly dependent on both the ability of the inhibitor and metal surface to form chemical bonds and the charges of the surface and inhibitor. Inhibitors can either be added to the service environment, or can be incorporated into coating systems, in which the coating typically contains a physical barrier layer that physically suppresses the corrosion process and an inhibitor layer that chemically suppresses it.

When selecting an inhibitor, the corrosion engineer should consider its long-range effectiveness, the magnitude of suppression of uniform and localized corrosion, its effect on bimetallic coupling to other metals joined to the main system, the effect of temperature and concentration on the performance of inhibitors, its effect on the existing condition of the system to be protected, its effect on heat transfer characteristics, possible toxicity and pollution problems, its economic and technical competitiveness with other considered inhibitors, etc.

In choosing a suitable corrosion inhibitor, it is important to know the corrosion potential with respect to the potential of zero charge, E_{pzc}. If $E_{corr} > E_{pzc}$, the excess surface charge density will be positive, and a negatively charged inhibitor may be the better choice. If E_{corr} occurs at a negative rational potential, a cationic inhibitor may be preferable. Neutral molecules can best serve as inhibitor if $E_{corr} \approx E_{pzc}$. It should be borne in mind that many of the commercial inhibitors are weak acids or bases, and their charge depends on pH. Thus, an inhibitor that acts well in a medium of low pH may be quite useless in a medium of high pH, and vice versa. When the usual aqueous medium is replaced by a nonaqueous or mixed solvent, the situation can change dramatically. The charge on the inhibitor molecule may be quite different, and E_{pzc} also depends on the solvent. The

solubility of an inhibitor used in aqueous solutions may also be quite different in a nonaqueous solvent. This changes the surface coverage corresponding to a given bulk concentration. To a first approximation, the surface coverage should be similar in different solvents, when comparison is made on a normalized scale of concentration, $c_b/c(\text{sat})$, obtained by dividing the concentration in each solvent by its saturation value. Thus, a different range of concentrations of inhibitor may have to be used for each solvent.

Inhibitors may be classified as (i) inorganic, (ii) organic, and (iii) environmental conditions. Since both anodic and cathodic reactions are involved in the corrosion process, corrosion inhibitors can decrease the corrosion rate if they can slow down any of these two reactions, or both. Inorganic inhibitors can be either anodic or cathodic.

Anodic inhibitors are compounds that suppress the anodic reaction. These include chromates, phosphates, nitrites, molybdates, borates, benzoates, dissolved oxygen, etc. Although chromate inhibitors have demonstrated the highest corrosion inhibitor performance and are versatile and cost-effective, they are toxic and harmful to the environment. They are widely used in cooling water, heat exchangers, automobile radiators, diesel and gas engines, power transformers, gasoline pipelines containing water, etc. The inhibitive properties of certain phosphates are based on their hydrolysis and increase of pH to above 8.0, and on formation of ferric and ferrous phosphate passive films on an iron surface. Phosphates protect carbon steel at temperatures up to 80 °C, but their efficiency is lower than at 20 °C. In hard water, the efficiency of phosphates is higher, thanks to the formation of mixed phosphates of calcium, magnesium, and iron on the protected metal surface. It should be noted that if a decrease in the exchange current density in the presence of an anodic inhibitor is due to inhibitor adsorption effectively decreasing the surface area, then if the anodic area is not completely covered, the cathode-to-anode surface area ratio will be increased to a high value, and severe pitting will occur at exposed anodes. For this reason, anodic inhibitors must be used with caution, and cathodic inhibitors are generally preferred.

Cathodic inhibitors decrease the rate of the cathodic reaction. They do not cause local attack if their concentration is less than the value needed for effective protection. Examples include zinc salts (e.g. $ZnSO_4$ or $ZnCl_2$), polyphosphates, poisons such as As, P, Sb, and Bi, and precipitators such as calcium bicarbonate ($Ca(HCO_3)_2$). If an inhibitor such as zinc is added to the corrosion system, it would react with the hydroxyl ion and precipitate insoluble compounds that would choke the cathodic sites on the metal surface. Poisons should be used with great caution as they might promote hydrogen-induced blistering or hydrogen sulfide cracking.

Organic inhibitors are film formers; they consist of amines, amine salts, esters, and ammonia derivatives.

Environmental conditions refer to scavengers such as sodium sulfite, ammonium sulfite ($(NH_4)_2SO_3$) and hydrazine (N_2H_4), and biocides such as ozone. Scavengers are chemicals that eliminate oxygen from closed systems.

Inhibitors can also be classified on the basis of their functions. For instance, chromates and nitrates are called *passivating inhibitors* because of their tendency to passivate the metal surface. Referring to Figure 17.17, we can say that line 1

represents a condition where there is no passivating inhibitor, line 2 represents a condition where there is insufficient passivating inhibitor, and line 3 represents a condition where there is sufficient passivating inhibitor. Some inhibitors, such as silicates, inhibit both the anodic and cathodic reactions. They also remove undesirable suspended particles from the system, such as iron particles, by precipitation.

Certain types of inhibitors make the surrounding environment alkaline to prevent corrosion. Such inhibitors in the gas phase are called "*volatile corrosion inhibitors*" (VCIs), or "vapor phase inhibitors" (VPIs), and they consist of heterocyclic compounds, such as cyclohexylamine. On contact with the metal surface, the vapor of these salts condenses and reacts with moisture to liberate protective ions. In boilers, volatile basic compounds, such as morpholine or hydrazine, are transported with steam to prevent corrosion in condenser tubes by neutralizing acidic carbon dioxide or by shifting surface pH toward less acidic and corrosive values. In closed vapor spaces, such as packing crates for sea transportation, volatile solids such as salts of dicyclohexylamine, cyclohexylamine, and hexamethylene-amine are used.

Migrating corrosion inhibitors (MCIs) were developed for reinforced concrete structures. They are classified as mixed inhibitors, meaning they affect both anodic and cathodic reactions. MCIs are both inorganic compounds (sodium-mono-fluoro-phosphate) and organic aqueous mixtures based on volatile amine constituents (e.g. amine alcohol) and nonvolatile ones, mainly an amino carboxylate compound or a phosphorus-containing compound. MCIs migration through the capillary structure of concrete is done first by liquid diffusion via the moisture that is normally present in concrete, next by its high vapor pressure, and finally through hairlines and microcracks. The effectiveness of different mechanisms depends on concrete cover, porosity, water content, solubility, and volatility of the migrating inhibitor. MCIs inhibit corrosion by penetrating concrete, ionically attracting to the rebar steel, and forming a protective molecular layer on the steel surface, thus hindering steel dissolution. Some researchers have also reported that MCIs help decrease the free chloride content in concrete and/or decrease the rate of diffusion. MCIs can be incorporated as an admixture, or can be surface impregnated on existing concrete structures. The migrating amine-based corrosion inhibiting admixture can be effective when incorporated in the repair process of concrete structure. Thus, this technology was used to rehabilitate the exterior Pentagon walls. The outcome of using MCIs is great extension of the concrete service life.

Using Evans diagrams, we can summarize some key effects of inhibitors on the values of the exchange current density and Tafel slopes, and consequently on the corrosion current density and corrosion potential. This is shown in Figure 17.31.

17.7.4 Coatings

Coating can be sorted into two groups: (i) active coatings such as zinc, which acts as a sacrificial anode even after part of the underlying metal has been exposed to the environment, and (ii) barrier coatings made of a more noble metal, such as nickel/chrome on steel or silver/gold on copper, or all sorts of paints. Such

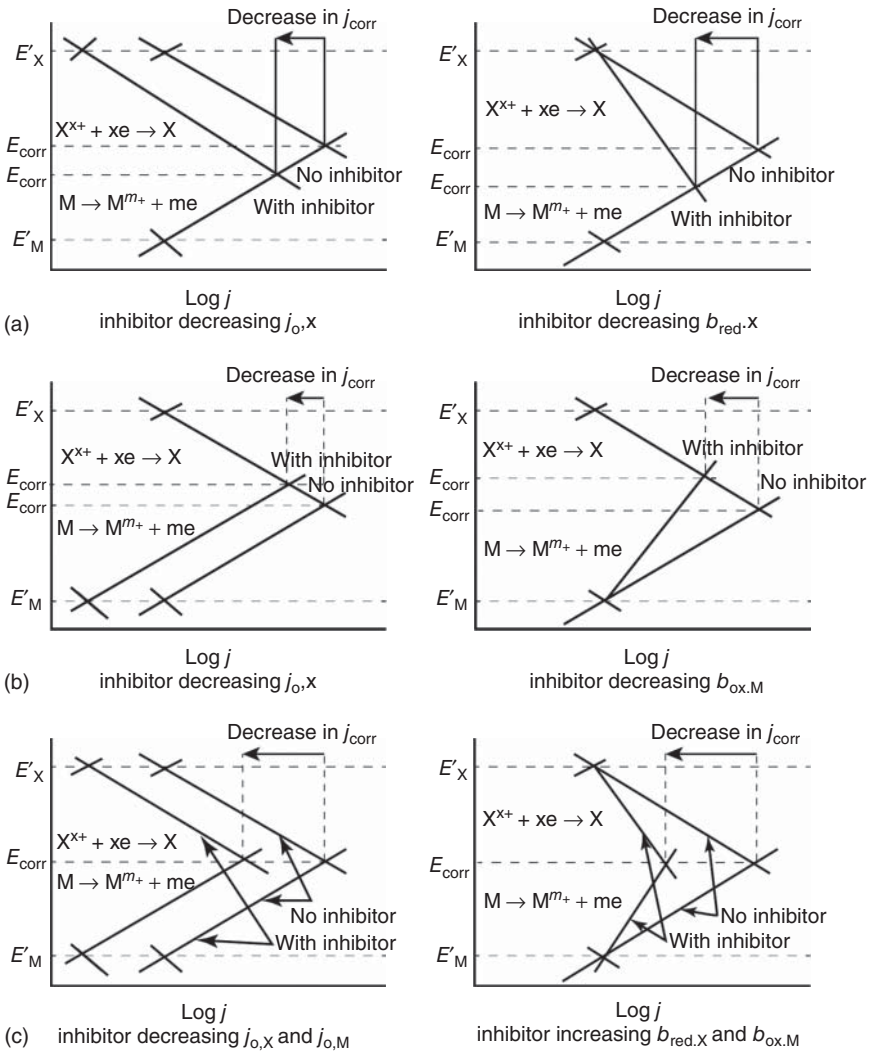


Figure 17.31 Evans diagrams illustrating some possible effects of corrosion inhibitors on the exchange current densities and Tafel slopes, and consequently on the corrosion current density and corrosion potential. (a) Cathodic inhibitor. Both E_{corr} and j_{corr} are decreased. (b) Anodic inhibitor. While E_{corr} is increased, j_{corr} is decreased. (c) Mixed (cathodic and anodic) inhibitor. While E_{corr} remains unchanged, j_{corr} is decreased. Source: Reprinted with permission from E. E. Stansbury and R. A. Buchanan, *Fundamentals of Electrochemical Corrosion*, ASM International, Materials Park, Ohio, 2000.

coatings prevent corrosion by simply isolating the metal from the environment. They can be excellent as long as they are intact. However, once damaged, galvanic corrosion (in the case of more noble metals) and differential aeration (in the case of nonmetallic coatings) may lead to an increase in the rate of pitting corrosion on the exposed areas. In this context, we should mention that surface preparation is a major factor in obtaining good, adherent coatings of any type. Degreasing,

chemical cleaning and, in some cases, mechanical treatment are essential steps in the preparation of the surface for coating, which often consists of several layers, for optimum protection.

The use of organic or metallic barrier coatings between a possibly aggressive environment and a material has become the most widely used method of protecting most engineered materials. There are three general classes of coatings: organic, inorganic, and metallic coatings. However, protective coatings often refer to integrated multifunctional systems that may combine more than one type of coating.

Organic coatings protect more metal on a weight basis than any other means of corrosion protection. Besides providing protection by creating a physical barrier between the metal and the environment, these coatings may also contain corrosion inhibitors or other additives to stifle corrosion processes. Organic coatings include paints, varnishes, resins, and lacquers. *Paint systems* comprise of the primer, an intermediate coat, and a top coat. The primer must be strongly bonded to the substrate, resistant to corrosive environments and suppress corrosion, and provide good adhesion to the intermediate coat or the top coat. *Varnish* is a preparation consisting of resinous matter dissolved in oil (oil varnish), alcohol (spirit varnish), or another volatile liquid. When applied to the surface of wood, metal, etc., it dries and leaves a hard, more or less glossy, usually transparent coating. Varieties of varnish include resin varnish, spar varnish, polyurethane varnish, acrylic varnish, and lacquer (quick-drying, solvent-based varnishes or paints). After application, the film-forming substances in varnishes may harden immediately, as soon as the solvent has fully evaporated. Otherwise, they may need to undergo certain curing processes, primarily chemical reactions between oils and oxygen, and chemical reactions between components of the varnish. Environmental factors such as heat and humidity play a very large role in the drying and curing times of varnishes. The drying and curing time of all varnishes may be shortened by exposure to an energy source such as sunlight, ultraviolet light, or heat. *Lacquers* are solutions of organic film-forming materials in organic solvents from which the solvent evaporates after application to a substrate. The solid film is formed after evaporation. The properties of lacquers depend on the type of resin used. Lacquers dry very rapidly. The major types of lacquers include cellulose derivatives, polyvinyl chloride (PVC), chlorinated rubber, and polyurethane elastomers.

Inorganic coatings include enamels, glass linings, and conversion coatings. Porcelain enamel coatings are inert in water and resistant to most weathers. They are common on appliances and plumbing fixtures. Glass-lined metals are used in process industries where there is concern over corrosion or contamination of the product.

When stainless steel products are manufactured, free iron is transferred to the surface of the material from the steel cutting, stamping, and forming tools used in the manufacturing process. Free iron can also be imparted on the surface by polishing or blasting operations that utilize the same polish or blast media between both mild steel and corrosion-resistant steel grades. Free iron readily oxidizes, forming visible rust on the surface of the product. Passivation of stainless steels is a chemical treatment with a specific acid formulation that removes free iron or

other surface contamination from the stainless steel while simultaneously promoting the formation of a passive chromium/nickel oxide layer to act as a barrier to further corrosion. When the surface iron is removed, the other components of the alloy (primarily chromium, often nickel as well) are left behind as a surface layer over the underlying steel. Upon exposure to air, they react with oxygen to form an oxide layer that protects the rest of the steel from corrosion. This corrosion-resistant surface can be damaged through mechanical means, heat, or chemical damage. When that happens, iron is exposed and the item is once again subject to rusting. For this reason, passivation may need to be performed on a regular basis. Passivation treatment usually consists of cleaning the stainless steel and then immersing it in a concentrated HNO_3 solution (in the case of stainless steels that contain more than 17 wt% Cr), or in a solution of HNO_3 with oxidizing salts (in the case of stainless steels for machining, polished surfaces, or stainless steels with Cr-content lower than 17 wt%). This process forms a transparent passive layer, typically about 3 nm thick. Applicable standards include ASTM A380, A967 and F86, AMS 2700, ISO 16048, and QQ-P-35 specifications. Passivation of high-strength steels should be followed by hydrogen release.

Conversion coatings are produced by intentionally corroding the metal surface in a controlled manner, in order to convert the substrate into a compound with desirable properties. This is done to produce an adherent corrosion product that protects the metal from further corrosion. Examples include anodizing, phosphatizing, black oxide, and chromate conversion coatings.

Anodizing, one of the most common conversion coatings, involves the electrolytic oxidation of typically aluminum, magnesium, titanium, or zirconium surfaces, to produce an oxide scale that is thicker and more adherent than the naturally occurring film. An anodic electrical current passing through an electrolytic bath converts the metal surface to a durable aluminum oxide. The difference between plating and anodizing is that the oxide coating is integral with the metal substrate, as opposed to being a metallic coating added by deposition. The most common electrolytes used are chromic acid, sulfuric acid, phosphoric acid, and their mixtures. Anodizing usually has three types: Type I – anodizing in chromic acid (1.3–7 μm thick layer is formed); Type II – anodizing in sulfuric acid (2.5–25 μm); and Type III – hard anodic coating (2.5–100 μm). It also has two classes: Class 1 – natural color; Class 2 – coating with paint. Anodizing is carried out after all machining, thermal treatments, and joining processes. Advantages include increase of corrosion resistance (in particular if the top porous layer is painted or sealed), increase in paint adhesion to the substrate, special decorative and optical properties, electrical insulation (dielectric), and abrasion resistance provided by the hard anodic coating. On the down side, this coating might decrease the fatigue strength (by as high as 30%).

In *phosphatizing*, the conversion coating consists of a thick porous layer of fine phosphate crystals, tightly bonded to steel, zinc, or aluminum. For many years, car bodies have been treated this way prior to coating. The steel is pickled or otherwise well cleaned and immediately held in a bath of hot phosphoric acid containing zinc and possibly manganese salts, in addition to some other additives. Phosphate coatings are porous; therefore, they form an attractive base for application of paints and organic coatings. The coating alone does not provide

significant corrosion resistance, but when combined with a top layer of paint, or a sealing film of oil or grease, the corrosion resistance is substantially increased. Phosphating should not be applied to nitrided or finish machined steel, and steel parts containing aluminum, magnesium, or zinc are subject to pitting in the bath. Some restrictions also apply to heat-treated stainless and high-strength steels.

Chromate conversion coatings are generally used on aluminum as well as on zinc and certain other metals. Chromate coatings are mostly applied on aluminum components, particularly on aircraft components, and on aluminum objects that are to be painted. They are also beneficial in marine environments. Chromium coatings provide a non-porous bonding surface for all paints. Chromium coating is formed by chemical attack on a metal when it is brought in contact with an aqueous solution that consists of a mixture of chromic acid, hydrofluoric acid, phosphoric acid, mineral acids, chromium salts, and some other salts. Spraying or application by brushing or swabbing can also be used for touchup of parts. Although chromate ions are known for their inhibition properties, concerns related to the toxicity of the hexavalent chromium used in chromating have limited its use in recent years.

Steels containing nitride-forming elements such as Cr, Mo, Al, and V can be treated to produce hard surface coatings with improved wear resistance. *Nitriding* typically involves exposure of cleaned surfaces to anhydrous ammonia at elevated temperatures. Because the nitrides establish compressive residual stresses on the treated surface, nitrided steels usually exhibit improved fatigue and CF resistance. Similar beneficial effects can be achieved by shot peening.

Metallic coatings and other inorganic coatings are produced using a variety of techniques, including hot dipping, electroplating, cladding, thermal spraying, chemical vapor deposition (CVD), or surface modification using directed energy (laser or ion) beams. The most widely used metallic coating method for corrosion protection is *galvanizing*, which involves the application of metallic zinc to carbon steel for corrosion control purposes. Hot-dip galvanizing is the most common process, and as the name implies, it consists of immersing iron or steel for 1–2 minutes in a bath of molten Zn at 445–465 °C. Another process is electrolytic galvanizing. Galvanizing can be found in almost every major application and industry where iron or mild steel is used. The simplicity of the galvanizing process is a distinct advantage over other methods of providing corrosion protection. The automotive industry depends heavily on this process for the production of many components used in car manufacturing, including fully assembled car bodies.

Several other coating processes are shortly mentioned here. *Thermal spraying* (sometimes referred to as *metallizing*, when depositing metals) techniques are coating processes in which melted (or heated) materials are sprayed onto a surface. The "feedstock" (coating precursor) is heated by electrical (plasma or arc) or chemical means (combustion flame). Thermal spraying can provide thick coatings (thickness may range from 20 μm to several millimeters, depending on the process and feedstock), over a large area at high deposition rate as compared to other coating processes such as electroplating, physical vapor deposition (PVD), and CVD. Coating materials available for thermal spraying include metals, alloys, ceramics, plastics, and composites. They are fed in powder or wire form, heated to a molten or semimolten state, and accelerated toward substrates in the form

of micrometer-sized particles. The coatings are formed by the accumulation of numerous sprayed particles. The surface may not heat up significantly, allowing the coating of flammable substances. Several variations of thermal spraying include plasma spraying, flame spraying, high-velocity oxy-fuel (HVOF) coating spraying, and cold spraying. One of the most common examples of thermal spraying for corrosion protection is arc spray deposition of Al–Zn–In alloy on the San Luis Pass Bridge for cathodic protection. *CVD* is a chemical process used to produce high-quality, high-performance, solid materials. In typical CVD, the substrate is exposed to one or more volatile precursors, which react and/or decompose on the substrate surface to produce the desired deposit. Frequently, volatile byproducts are also produced, which are removed by gas flow through the reaction chamber. Thick (several mm), dense coatings can be obtained by CVD. This process is common when the melting temperature of the deposited material is high. *Diffusion coating* refers to a process where the surface of the metal to be coated is modified by diffusing into it at a high temperature into a metal or an element, which would provide the required resistance when combined with the parent metal. Diffusion coatings can be applied to a range of metals and alloys, such as Ni, Ti, and Mo, but the widest use is on ferrous metals. Examples are Zn or Al diffusion coatings.

Metal layers of varying thicknesses can be applied to other metals by various methods designed to obtain a more corrosion-resistant surface while using a less expensive or a stronger structural material to fabricate the component. *Cladding* is most commonly applied at the mill stage by the manufacturers of sheet, plate, tubing, or coins. Cladding by pressing, rolling, or extrusion can produce a coating in which the thickness and distribution can be controlled over wide ranges, and the coatings produced are free of porosity. The bonding of aluminum to steel is obtained by rolling at 540 °C. Alclad is a trademark name of Alcoa for wrought aluminum products, such as sheet and wire, clad with a layer (approximately 5.5% thickness per side) of high-purity aluminum. It is popularly used in airplane manufacture. The clad (coating) is anodic to the core; thus, it provides cathodic protection to exposed areas on the core.

17.7.5 Other Mitigation Practices

Other mitigation practices, beyond cathodic and anodic protection, inhibitors, and coatings, have been described in Section 17.6 for each corrosion form, and will not be described here again. For example, stirring velocity affects the concentration polarization regime. In the case of a metal with a passivation layer, above a certain stirring velocity there is no longer effect on the corrosion rate. In contrast, in the case of a metal without passivation layer, increased stirring velocity increases the corrosion rate within the active regime. At high stirring velocity, the corrosion rate attains a minimum value due to entrance into the passivity regime. In the design stage, drains should always be located at the lowest point in a tank, and the joint between the tank and the drain pipe should be designed to permit free drainage. It is good practice to support small to moderate size tanks on stanchions rather than on pads. Metals sitting on concrete pose a special problem

because the concrete is porous and holds moisture that eventually might cause crevice corrosion of the metal.

Further Reading

1 Ifergane, S., Ben David, R., Sabatani, E. et al. (2018). Hydrogen diffusivity and trapping in Custom 465 stainless steel. *J. Electrochem. Soc.* 165: C107–C115.
2 Popov, B.N. (2015). *Corrosion Engineering: Principles and Solved Problems*. Amsterdam, The Netherlands: Elsevier.
3 Hihara, L.H., Adler, R.P.I., and Latanision, R.M. (ed.) (2014). *Environmental Degradation of Advanced and Traditional Engineering Materials*. Boca Raton, FL: CRC Press.
4 Revie, R.W. (2011). *Uhlig's Corrosion Handbook*, 3rd edn. New York: Wiley.
5 Groysman, A. (2010). *Corrosion for Everybody*. Dordrecht, Heidelberg: Springer Science+Business Media.
6 Roberge, P.R. (2008). *Corrosion Engineering: Principles and Practice*. New York: McGraw-Hill.
7 Riskin, J. (2008). *Electrocorrosion and Protection of Metals: General Approach with Practical Consideration to Electrochemical Plants*. Oxford, UK: Elsevier.
8 Little, B.J. and Lee, J.S. (2007). *Microbiologically Influenced Corrosion*. New Jersey: Wiley.
9 Marcus, P. and Mansfeld, F. (ed.) (2006). *Analytical Methods in Corrosion Science and Engineering*. Boca Raton, FL: CRC Press.
10 Mansfeld, F. (2006). Classic paper in corrosion science and engineering with a perspective by F. Mansfeld. *Corrosion* 62: 843–855.
11 Baboian, R. (ed.) (2005). *Corrosion Tests and Standards – Application and Interpretation*, 2nd edn. PA: ASTM International.
12 Giriga, S., Kamachi Mudali, U., Raju, V.R., and Baldev, R. (2005). Electrochemical noise techniques for corrosion assessment – a review. *Corros. Rev.* 23: 107–170.
13 Kelly, R.G., Scully, J.R., Shoesmith, D.W., and Buchheit, R.G. (2003). *Electrochemical Techniques in Corrosion Science and Engineering*. New York: Marcel Dekker.
14 Baboian, R. (ed.) (2002). *NACE Corrosion Engineer's Reference Book*. Houston, TX: NACE International.
15 Hong, S.-B., Eliaz, N., Sachs, E.M. et al. (2001). Corrosion behavior of advanced Ti-based alloys made by three-dimensional printing (3DP™) for biomedical applications. *Corros. Sci.* 43: 1781–1791.
16 Davis, J.R. (ed.) (2000). *Corrosion: Understanding the Basics*. Materials Park, OH: ASM International.
17 Jones, D.A. (1995). *Principles and Prevention of Corrosion*, 2nd edn. Prentice-Hall.
18 Turnbull, A. (1994). Standardisation of hydrogen permeation measurement by the electrochemical technique. In: *Hydrogen Transport and Cracking in Metals* (ed. A. Turnbull), 129–141. UK.

19 Fontana, M.G. (1986). *Corrosion Engineering*, 3rd edn. Singapore: McGraw-Hill.
20 Pourbaix, M. (1974). *Atlas of Electrochemical Equilibria in Aqueous Solutions*, 2nd edn. Houston, TX: NACE.
21 Pourbaix, M. (1973). *Lectures on Electrochemical Corrosion*. New York: Plenum Press.
22 Wagner, V.C. and Traud, W. (1938). Über die deutung von korrosionsvorgängen durch überlagerung von elektrochemischen teilvorgängen und über die potentialbildung an mischelektroden. *Z. Elektrochem. Angew. Phys. Chem.* 44: 391–402.
23 *Corrosion*. http://www.corrosionjournal.org/.
24 *Corrosion Reviews*. https://www.degruyter.com/view/j/corrrev.
25 *Corrosion Science*. http://www.sciencedirect.com/science/journal/0010938X/.
26 *Corrosion and Materials Degradation*. http://www.mdpi.com/journal/cmd.
27 *Materials Degradation*. https://www.nature.com/npjmatdeg/.

18

Electrochemical Deposition

18.1 Electroplating

18.1.1 Introduction

Electrodeposition may be defined in a broad manner as the process of depositing a substance upon an electrode by electrolysis. Electroplating is conducted by passing an electrical current through a solution containing dissolved metal ions and the (usually metal) object to be coated. The metal object serves as the cathode in an electrochemical cell. The process is regulated by controlling a variety of parameters, including the voltage/amperage, temperature, bath chemistry, pH, etc. The process specification may include a sequence of steps, such as preplate machining of the substrate, stress relief, alkaline and acidic cleaning steps, stripping of old plating or paint, electroplating, rinsing in cold or hot water between steps, and postplate baking to remove hydrogen. Electroless plating uses similar steps but involves the deposition of metals (and other materials) on a substrate without the application of external electrical energy. In this chapter we shall discuss electroplating and electroless plating, as well as electrophoretic deposition (EPD). Underpotential deposition (UPD) was the subject of Chapter 11, and thus will not be discussed in this chapter again.

Electrodeposition has been practiced in industry for more than 150 years. For many years it was considered as an empirical low-tech technology. Although many useful plating baths were developed and additives were identified for different purposes – macro and micro leveling, brightening, stress relieve, increase of current efficiency (also known as faradaic efficiency, FE), inhibition of hydrogen embrittlement, and so on – most of the progress was made by ingenious trial-and-error methods, with relatively little effort toward determining the mechanism involved.

A turning point in this approach can be associated with the replacement of chemical vapor deposition (CVD) of aluminum by electroplated copper wiring in ultra-large-scale integration (ULSI) silicon chips, announced by IBM in 1997. This led to increased interest in electroplating in the microelectronics industry and to awareness of possible advancements through research and development, both in industry and in academia. The introduction of copper plating as an integral part of manufacturing microprocessors was no minor feat and came after about a decade of intense research and development. Following its success,

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

electroplating was elevated from the status of an empirical technology to that of a high technology, based on research and profound understanding of the way the different factors influence the quality of the product.

Electroplating offers several important advantages compared to most other plating technologies: (i) it is a relatively low-cost process; (ii) fairly simple and available equipment is needed; (iii) the laws governing scaling up and scaling down of electrochemical processes are well understood; (iv) porous, geometrically complex, or non-line-of-sight surfaces can be coated; (v) proper design of the cell and the counter electrode can ensure that metal is deposited only where it is needed; (vi) the throwing power is higher than in physical vapor deposition (PVD); (vii) high-aspect-ratio structures with good precision are produced; (viii) the relatively low processing temperature allows the formation of highly crystalline deposits, with possibly lower residual stresses; (ix) the thickness, composition, and microstructure of the deposit can be controlled precisely; (x) dense materials with high purity, low defect density, and narrow distribution of grain size can be produced.

Representative applications of electrodeposition include, among others, Au-plated brass jewelry; Cu plating for fabrication of interconnects in electronic packaging; hard Cr plating of aircraft landing gears made of alloy steels; Ag-plated mirrors; plating of Ni, Ni–Fe, Cu, and Pt–Co alloys in fabrication of micro-electromechanical systems (MEMS) by LIGA; W- and Re-based coatings for high-temperature and wear resistance, etc.

18.1.2 The Fundamental Equations of Electroplating

The fundamental equation of the overall reaction of metal deposition is

$$M^{z+} + ze^- \rightarrow M \tag{18.1}$$

where M is any metal with a valence z. But, this equation does not specify the phase in which each species is located and, more importantly, it does not take into account the role of the solvent. Both can be accounted for by rewriting the equation in the form

$$[M(H_2O)_n]^{z+} + ze_M^- \rightarrow M_M^0 + n(H_2O) \tag{18.2}$$

where the number of water molecules in the solvation shell need not be specified, because it may be different for different metals. It is important to realize that the hydration energy of ions in solution, U_{hyd}, is very high, of the order of

$$U_{hyd} = 5 \times z^2 \ \text{(eV)} \tag{18.3}$$

It is not implied here that all ions having a given valence possess the same energy of hydration. For any given valence, smaller ions have higher energies of hydration, but the factor of z^2 does show that valence is the main factor determining this energy.

Now, in order to deposit a metal ion from solution, it is necessary to remove all its solvation shell, which requires a lot of energy. This is not a thermodynamic barrier since much if this energy is regained in the overall process. However, the high energy of hydration is expected to play a major role in the kinetics of the

deposition process, because this energy, or at least part of it, has to be supplied in order to allow the metal ion to reach the surface and interact with it. Thus, one would expect that metal deposition processes would be very slow, compared to outer-sphere charge-transfer processes. This is not borne out by experiment. Indeed, the exchange current densities observed for metal deposition are often as high, or even higher, than outer-sphere charge-transfer processes. A commonly accepted mechanism explaining this unexpected behavior has yet to be developed.

18.1.3 Practical Aspects of Metal Deposition

In the current chapter, the emphasis is not on the mechanistic aspects of metal deposition, but rather on its practical aspects, such as uniformity of plating, the solution chemistry, and corrosion resistance. Diffusion of adatoms on the surface, from their initial landing site to edges, kinks, or vacancies on the surface, is considered to be the rate-determining step for metal deposition in some cases. Impurities or additives adsorbed on the surface can hinder such diffusion and can control the surface morphology of the resulting deposit.

It is interesting to consider the metal deposition process from a microscopic point of view. A rate of 20 mA cm^{-2} corresponds to the deposition of about 50 atomic layers of metal atoms per square centimeter per second. This may be too fast for the adatoms to reach their equilibrium positions. It is indeed observed that the alloys formed during electrodeposition are not necessarily those corresponding to the phase diagram of the metals involved.

The interaction between the substrate and the metal being deposited can also play an important role in determining the quality of a plated product. When the crystal parameters of the two metals are different, one of two situations may be observed: (i) The metal being plated may initially attain the crystal structure of the substrate, although this is not its most stable form. This is referred to as *epitaxial growth*. As the thickness of the deposit grows, it gradually reverts to its stable crystal structure. The stress created by the epitaxial growth can be relaxed by impurity atoms and by dislocations in the metal. (ii) Large differences between the crystal structures of the two metals do not favor epitaxial growth. In such cases, a so-called *crystallization overpotential* is observed (see Chapter 5), followed by two-dimensional (2D) nucleation on the surface. The effect is similar to the formation of small crystals in a supersaturated solution, or of droplets in a vapor.[1]

The current efficiency is the fraction of the total current used to deposit the metal (or metals, during alloy deposition). Side reactions, mostly hydrogen evolution, play an important role in electroplating. As a rule, their effect is detrimental to the process, because of the loss of energy and possible hydrogen embrittlement (HE).

Finally, there is the question of uniformity. Parts to be plated are rarely flat. They have grooves, edges, corners, protrusions, and so on. A good plating bath, which covers a surface uniformly irrespective of its shape, is said to have a good

1 Note that a 10-fold supersaturation is equivalent, in terms of Gibbs energy, to an overpotential of only 29.5 mV for deposition of a divalent metal.

throwing power. The factors controlling the throwing power of plating baths are discussed in detail in Sections 18.1.6 and 18.1.7 where we discuss some of the practical aspects of electroplating.

18.1.4 Hydrogen Evolution as a Side Reaction

Plating from aqueous solutions is limited by the competing reaction of hydrogen evolution. Considering the standard thermodynamic potentials of metal, it is noted that only Ag ($E^0 = +0.799$ V) and Cu ($E^0 = +0.337$ V) have positive potentials compared to the SHE (see Table 17.1), and therefore can be deposited even from acid solutions with 100% current efficiency. Pb (−0.126 V) and Sn (−0.136 V) can also be deposited readily, because the exchange current density for hydrogen evolution on these metals is very small, and the rate of hydrogen evolution at potentials where these metals are deposited is negligible. Proceeding to more active metals (i.e. those having a more negative standard potential), such as Cd (−0.403 V), Zn (−0.763 V), and Cr (−0.744 V), electroplating is still possible, but the current efficiency is lower. This leads, of course, to a loss of energy, but that is usually a minor consideration. Copious formation of hydrogen bubbles may screen parts of the surface from the solution, leading to nonuniform coating. In addition, a safety hazard may exist, unless the plating shop is very well ventilated.

The worst aspect of hydrogen evolution during metal deposition is the formation of atomic hydrogen on the surface, as an intermediate. A fraction of these atoms can diffuse into the metal, potentially causing HE and catastrophic failure (see Chapter 17). In coating systems with low FE, hydrogen evolution during electrodeposition is enhanced, thus increasing the likelihood of HE. For example, hard chromium (FE ~ 15%)-coated items are more prone to HE than those coated with cadmium. In highly efficient systems, such as Cu and Ag baths, codeposition of hydrogen occurs only when the limiting current density is exceeded, or when the added complexing agents shift the potential of metal deposition to sufficiently negative values. The higher the hydrogen overpotential on a given metal, the lower the amount of hydrogen absorbed in it. Heat treatment ("baking") is commonly employed following electroplating of various coatings in order to render the normally mobile hydrogen immobile. For high-strength steels electrodeposited with hard chromium, for example, this is typically done at 177–205 °C for at least three hours (according to the QQ-C-320 standard). The treatment is required for all steel parts hardened to 40 HRC or above, and should be applied not later than four hours after the completion of the plating process. Evaluation of the effect of residual hydrogen on the mechanical properties is then done according to standards such as ASTM F519 and ASTM F1624. Certain very susceptible alloys require the use of fabrication processes other than electroplating. Although hydrogen diffusion is exponentially increased with temperature, the required time to reduce the diffusible hydrogen concentration to a given level increases with the square of the section thickness. For thick sections this can mean hundreds of hours at any common baking temperature. Even then, there is no guarantee that permanent damage or irreversible HE has not already occurred. In the case of Cd plating, for example, the high solubility of hydrogen in Cd

relative to that in the iron-based substrate makes the Cd-layer a source during the initial stage of baking; consequently, the hydrogen concentration in the steel substrate may increase. Moreover, because the diffusivity of hydrogen in Cd is much lower than that in steel, the Cd-layer acts as a diffusion barrier to outgassing of hydrogen during baking, and a significant concentration of dissolved hydrogen might remain in the steel, even after baking times of 100 hours. If standard procedures are followed precisely, HE-related failures of electroplated items can usually be prevented. However, if improper baking process was carried out, or such a treatment was not carried out at all, delayed failure might occur. Eliaz et al. (*Corros. Sci.*, **50** (2008) 1481) developed a sensitive analytical procedure based on SIMS to allow in-service identification of local hydrogen accumulation, either during quality control or during failure analysis of electroplated items. The main effect of baking was found to be effusion of hydrogen from the substrate steel/Cd coating interface and the substrate steel into the atmosphere. In non-baked samples, an increase in the hydrogen signal was found at the Cd/steel interface. The effect was monitored even after 16 months storage in a desiccator. Based on their experimental data, Eliaz et al. explained HE-related delayed failures in terms of the time-independent reservoir of hydrogen at the coating/substrate interface, rather than in terms of irreversible damage that occurred within the substrate during electroplating.

In the case of even more active metals, such as Al and Mg ($E^0 = -1.662$ and -2.363 V versus SHE, respectively), the side reactions become so dominant that metal deposition no longer occurs from aqueous solutions.

18.1.5 Plating of Noble Metals

Silver and copper, with their standard potential positive with respect to the SHE, are about the easiest metals to plate because, in the absence of ligands forming a complex with their ions, the deposition potentials are positive with respect to the reversible hydrogen electrode (RHE), and hydrogen evolution would not take place, even in acid solutions. Based on this argument, it would be expected that the noble platinum-group metals (namely, Ru, Rh, Pd, Os, Ir, and Pt) could also be electroplated readily. This is not the case, because these metals do not tend to form a stable solution of the simple hydrated ions, and most plating solutions contain a complex with a suitable ligand, lowering the standard potential significantly.

Choosing a suitable ligand often presents a challenge because the complex has to fulfill two inherently opposing requirements. On the one hand, the ligand should bond to the noble metal cation strongly enough to ensure stability of the plating bath. On the other hand, it should bond weakly enough to allow deposition of the noble metal. In the case of Pt, the two common complexes employed are $[PtCl_4]^{2+}$ and $[Pt(NO_2)_2(NH_3)_2]^0$. Interestingly, fresh solutions of either of these complexes do not perform well. In the case of the chloro-complex, aged solutions perform much better than fresh ones. On the other hand, in the case of the nitroso-amino complex, aging is not enough. The performance of a freshly prepared solution is improved with subsequent plating operations. Recent studies indicate that the inherently conflicting requirements of the ligand

have been solved in an elegant way. The freshly prepared solution is good for the stability of the bath, but is too stable for the plating operation itself. However, aging or passage of charge in the solution causes the replacement of one or more of the ligands by water, forming a less stable complex that allows efficient plating of the noble metal. Even so, deposition occurs at rather negative current densities, where hydrogen evolution takes place, in spite of the fact that the standard potential for Pt deposition is +1.200 V versus SHE.

18.1.6 Current Distribution in Electroplating

Current distribution on the cathode is a major variable in electrodeposition. It is determined by several parameters, such as the geometry of both the anode and the cathode, as well as their relative position in the bath, the kinetics of charge transfer, and the conditions of mass transport. It determines the thickness of the coating and its uniformity, as well as the local chemical composition in alloy deposition.

18.1.6.1 Uniformity of Current Distribution

A plating bath should produce a uniform thickness of plating. If metal deposition is the only reaction taking place, variations in the thickness of the deposit are an expression of the uniformity of current distribution pertaining to the specific conditions of plating. This, in turn, depends on the composition of the bath, which determines the conductivity in solution and has an influence on the kinetics of deposition. It also depends on the applied current density, the temperature, and the geometry of the part being plated. Edge effects can be troublesome, and the position of the counter electrode(s) with respect to the working electrode (which is the part being plated) can play a major role, sometimes overcoming the inherent limitations of the plating bath.

Nonuniformity of the current density at the edge of the substrate being plated was discussed in Section 3.2. The current distribution was discussed there for three simple geometries. In the real world, the geometry is rarely simple, and the current distribution cannot be calculated using the analytical solution of an equation. It can, however, be determined by numerical calculation using, for example, the finite element method (FEM). Commercial software is available to simulate different positions and shapes of the counter electrode with respect to the object being plated, eliminating the need for trial-and-error methods of obtaining the highest degree of uniformity. Incidentally, the same kind of software can be used to create any specific nonuniformity of coating thickness that may be required for specific applications. The expected uniformity of current distribution of a given plating bath can be estimated from its conductivity and kinetics, as discussed in the next section.

18.1.6.2 The Faradaic Resistance (R_F) and the Solution Resistance (R_S)

The resistance between the working electrode being plated (the cathode) and the counter electrodes (the anodes) can be represented by two resistors in series, as shown in Figure 18.1.

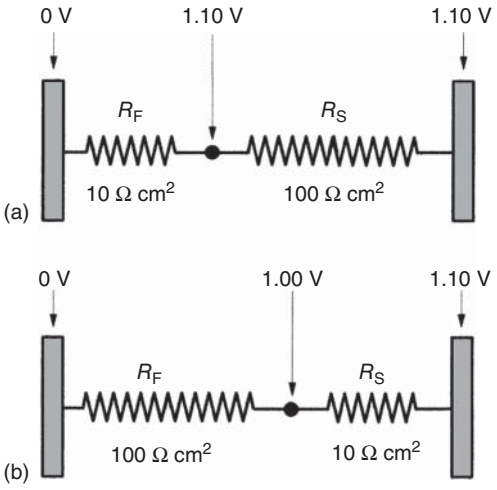


Figure 18.1 The values of R_F and R_S and the corresponding change in potential between the working and the counter electrodes at a plating rate of 10 mA cm^{-2}. (a) Primary current distribution, $Wa = 0.1$. (b) Secondary current distribution, $Wa = 10$. Wa is the dimensionless Wagner number; see Section 18.1.6.3.

The Ohmic solution resistance, R_S, is determined by the specific resistivity of the solution and by the configuration of the cathode and the anode with respect to each other, as well as their respective shapes. When this is the largest resistance in the system, the process is said to occur under conditions of *primary current distribution*. The relevant scale to be considered is of the order of a few millimeters up to several centimeters. Primary current distribution usually leads to nonuniform current distribution on the cathode, resulting in nonuniform thickness of the coating. It is hardly necessary to point out that nonuniform current distribution can lead to poor performance in any industrial process. For alloy deposition this may be even more critical, considering that the composition of the alloy is often a function of current density. Thus, a nonuniform current distribution might lead to nonuniformity of the alloy composition, in addition to variation of the thickness of the coating.

When the faradaic resistance, R_F, is the largest resistance in the system, the reaction is said to be "activation controlled," and *secondary current distribution* is maintained. It should be obvious that secondary current distribution leads to uniform thickness of the deposit, on the macroscopic scale. For example, in a barrel plating process for plating screws, primary current distribution will lead to excessive coating on the top of the grooves and no coating at the bottom. In contrast, secondary current distribution will lead to much more uniform coating thickness throughout the groove.

18.1.6.3 The Dimensionless Wagner Number

In considering Figure 18.1 it should be borne in mind that R_F is independent of the geometry of the cell, because the potential drop across this resistance occurs over a distance of less than 1 nm, while the solution resistance extends over the typical macro-dimensions of a plating bath, which is of the order of a few centimeters, namely, about 10^7 nm. Thus, for the case shown in Figure 18.1a, the potential drop over R_F is only 0.10 V, while that over R_S is 1.0 V. This represents the primary current distribution. Since R_S depends on the local geometry, this

will lead to nonuniformity of current distribution. In contrast, Figure 18.1b represents secondary current distribution. Most of the potential drop will occur across the faradaic resistance, and changes in geometry have little effect on the total resistance between the anode and the cathode, leading to nearly uniform film thickness.

The Wagner number is a dimensionless parameter that helps us determine the probable level of uniformity of current distribution in electroplating. There are two limiting cases to be discussed:

(a) *Primary current distribution*, where the uniformity is determined exclusively by the conductivity of the solution and the geometry of the cell.
(b) *Secondary current distribution*, where the current distribution is determined by the kinetic parameters of the deposition process. As usual, reality is somewhere in between; it is neither purely primary nor purely secondary. The transition from one regime to another is characterized by the *Wagner number*, defined as

$$Wa = \frac{(\partial \eta / \partial j)_{c_i,T}}{\rho L} \tag{18.4}$$

where ρ is the specific resistivity of the solution (in units of Ω cm) and L is the characteristic length.

The choice of the characteristic length L is not always obvious. It may be the length of the electrode being plated, the distance between the working and the counter electrodes, or "the dimension of the irregularity," for example, the difference between the shortest and the longest distance between the two electrodes. The last is the best choice in most cases, since it reflects the differences in the solution resistance on different areas of the item being plated. The exact value taken is not critical, however, since *Wa* should be used only as a guideline, and the actual current distribution (or variation of thickness of plating) can be found either experimentally or by obtaining a numerical solution for the specific geometry and the kinetic parameters of the metal deposition considered.

The partial derivative, taken at constant concentration (and, of course, constant temperature and pressure), is the differential faradaic resistance, R_F, in units of Ohm square centimeter. The solution resistance, expressed in the same units, can be written as follows:

$$R_S = \rho L \tag{18.5}$$

Hence, the Wagner number can be expressed simply as

$$Wa = \frac{R_F}{R_S} \tag{18.6}$$

In the absence of mass-transport limitations, the local current density at a given potential is determined by the sum of two resistors in series – the faradaic resistance and the solution resistance. For values of $Wa \ll 1$, the solution resistance is dominant, and the current distribution depends primarily on geometry. This is

the realm of primary current distribution. For $Wa \gg 1$, the faradaic resistance is predominant, and secondary current distribution is observed. What this means in the practical sense is that, under conditions of secondary current distribution, the resistance between the part being plated and the counter electrode is determined primarily by the resistance of the interphase to charge transfer; changes in the distance between the electrodes play a smaller role. In practice, a system is said to be under primary current distribution if $Wa \leq 1$, and under secondary current distribution if $Wa \geq 10$.

It is interesting to consider the value of the faradaic resistance at different overpotentials. Close to the reversible potential, the faradaic resistance is given by

$$R_F = \frac{\eta}{j} = \frac{1}{j_0}\frac{RT}{nF} \tag{18.7}$$

It is independent of potential and of the applied current density, but inversely proportional to the exchange current density, because the current–potential relationship is linear in this region (c.f. Section 5.2.4). In this region, the Wagner number is also inversely proportional to the heterogeneous rate constant of the metal deposition reaction. Thus, fast reactions have low value of the Wagner number and tend to lead to primary current distribution. This is a rather unique situation in electrochemistry, where poor catalytic activity (i.e. low specific rate constant) is an advantage.

Considering the high-overpotential region, the relevant rate equation is

$$j = j_0 \exp\left(-\frac{\alpha_c \eta F}{RT}\right) \tag{18.8}$$

It follows that the faradaic resistance in this region is given by

$$\frac{1}{R_F} = \left(\frac{\partial j}{\partial \eta}\right)_{c_i} = \frac{\alpha_c F}{RT} j \tag{18.9}$$

$$R_F = -\frac{1}{j}\frac{RT}{\alpha_c F} \tag{18.10}$$

It is concluded that at high overpotentials, the faradaic resistance is inversely proportional to the applied current, but independent of the exchange current density. Supposing that one compares two electrodeposition reactions having values of $j_0 = 10^{-4}$ A cm^{-2} and $j_0 = 10^{-6}$ A cm^{-2}, respectively, and deposition is conducted at a current density of $j = 10^{-2}$ A cm^{-2}. Both reactions will be in the linear Tafel region, and their faradaic resistance will be equal, as long as their Tafel slopes are equal. This is the reason for the often-used practice of plating initially at a low current density, to increase the Wagner number and, thus achieve better uniformity of current distribution, at least for a thin layer of coating in recessed areas. Subsequently, plating is done at a higher current density, to keep the time needed to reach the desired thickness reasonable.

The dependence of the Wagner number on the current density applied, for different values of the exchange current density, is shown in Figure 18.2. The region marked for $-\log j = 1.5$–2.5 (corresponding, approximately 3.0–30 mA cm^{-2})

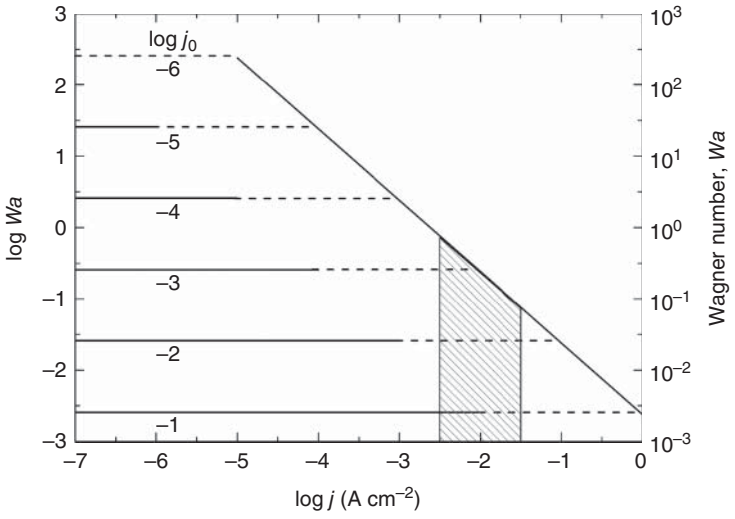


Figure 18.2 The Wagner number, calculated for different values of the exchange current density, as a function of the applied current density. $L = 2$ cm, $b = 0.12$ V, $\rho = 25\ \Omega$ cm, $n = 2$.

is typical for many electroplating processes and, therefore, warrants some discussion.

At low current densities, where $j/j_0 \leq 0.1$, the faradaic resistance is independent of the current density (see Eq. (18.5)), but is proportional to $1/j_0$. At high current densities, where $j/j_0 \geq 10$, it is independent of j_0 (see Eq. (18.7)), but is proportional to $1/j$. Needless to say, the transition from one region to another is smooth, not sharp as might be implied in this figure. Consider a deposition current density of 10 mA cm^{-2}. On extrapolation, the two straight lines in Figure 18.2 cross for a value of $j_0 = 10^{-3}$ A cm^{-2} at $Wa \approx 0.25$. For values of $j_0 \leq 10^{-4}$ A cm^{-2}, the system will be in the region where the Wagner number is independent of j_0. At higher values of the exchange current density, where $j_0 \geq 10^{-2}$ A cm^{-2}, it will be inversely proportional to j_0.

This leads us to the discussion of the methods by which the throwing power can be increased. The specific resistivity of the solution, ρ, is important. It could be decreased by adding a supporting electrolyte, but this approach is limited in scope. A value of $\rho \approx 5\ \Omega$ cm, measured in acid copper baths containing $CuSO_4$ and H_2SO_4, is about as low as one can go. Another approach is to decrease j_0 by adding a complexing agent, as discussed for Cu deposition from a pyrophosphate salt (see Section 18.1.6.4). It should be borne in mind, however, that when a metal ion is complexed, its standard potential is shifted cathodically by $(2.3RT/nF)\log K$, where K is the stability constant of the complex formed. As a result, hydrogen evolution can occur along with metal deposition, and the current (faradaic) efficiency might be decreased significantly.

A decrease in current efficiency with increasing current density is observed in many cases. This can enhance the uniformity of thickness of the deposit due to a negative feedback effect. Thus, in areas on the surface where the total current density is higher, the fraction of the current consumed for metal deposition is

lower. The dependence of the current efficiency on current density is governed by the kinetic parameters of the two reactions involved. Thus, in the commonly encountered situation where hydrogen evolution is activation controlled while metal deposition is partially controlled by mass transport ($0.05 \leq j/j_L \leq 0.7$), the current efficiency will decrease with increasing current density.

Electrodeposition of a metal from a negatively charged complex ion can influence the throwing power and the morphology of the deposit in other ways as well. Where the local current density is higher, the potential on the solution side of the interphase is more negative. This causes a decrease in the local concentration of the negative ions, which slows the reaction. In other words, a negative feedback mechanism is created, counteracting the variation of local current density caused by the primary current distribution.

18.1.6.4 Kinetically Limited Current Density

The case of mass-transport limitation has been discussed before. This leads to a mass-transport-limited current density, independent of potential. In practice, deposition of metals is conducted at current densities well below the limiting current, in the range of $j \leq 0.3j_L$. At higher current densities, the deposits tend to be rough, powdery, or friable.

However, a limited current density could also be observed when the limitation is kinetic. This can happen when the metal ion exists in solution as a complex, although the electroactive species participating in the charge-transfer step is the free ion. An example of such a situation is the deposition of Cu from a solution containing copper pyrophosphate. A chemical step, the rate of which is independent of potential, is followed by an electrochemical step, as shown by the next two equations:

$$Cu_2P_2O_7 \xrightarrow{k(\text{chem})} 2Cu^{2+} + P_2O_7^{4-} \tag{18.11}$$

$$2Cu^{2+} + 4e_M^- \rightarrow Cu_M^0 \tag{18.12}$$

A limiting current is observed for this reaction, but its value is between 1 and 2 orders of magnitude lower than that calculated for mass-transport limitation, indicating that it is due to the rate of release on Cu^{2+} ions from the complex. Now, the faradaic resistance, defined as the partial derivative $\partial\eta/\partial\phi$, approaches infinity at the limiting current, and so does the Wagner number, leading to full secondary current distribution and, hence, to uniform current distribution.

Indeed, this may be the reason for enhanced uniformity when plating is conducted from a solution containing a suitable ligand to form a stable complex in solution. Even in a less extreme case, when the chemically controlled limiting current, $j_{L,\text{chem}}$, is comparable to or even larger than the mass-transport-limited current density, j_L, improvement of the uniformity of deposit thickness may be achieved, because the two limiting currents add up as

$$\frac{1}{j} = \frac{1}{j_L} + \frac{1}{j_{L,\text{chem}}} \tag{18.13}$$

so that kinetic limitation could increase the value of the Wagner number.

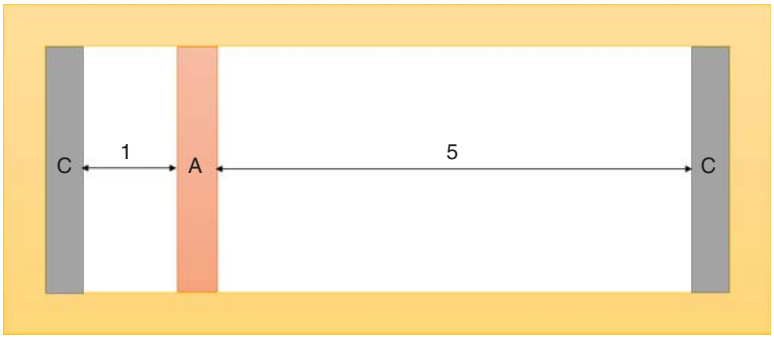


Figure 18.3 Top view of the Haring–Blum cell for the determination of the *macro throwing power*. A, anode; C, cathode; numbers reflect the ratio between the two cathode to anode distances. Source: H.E. Haring and W. Blum, *Trans. Am. Electrochem. Soc.*, **44** (1923) 313.

18.1.7 Throwing Power

18.1.7.1 Macro Throwing Power

The throwing power of a bath is a measure of its ability to produce electroplated coatings of uniform thickness on samples having complex geometries. A quantitative measure of this property can be obtained by employing the *Haring–Blum cell* shown in Figure 18.3. In this cell, two parallel cathodes connected electrically are positioned at unequal distances from two sides of an anode. The throwing power (TP) is defined as

$$\mathrm{TP} = \frac{K - M}{K} \times 100 \tag{18.14}$$

where K is the ratio of distances between the anode and the two cathodes (usually taken as 5) and M is the ratio of coating thicknesses on the two cathodes. This is a little awkward, because ideal throwing power, defined as uniform plating thickness irrespective of geometry, corresponds to $M = 1$, yielding a value of $\mathrm{TP} = 80\%$, rather than 100%, which one would expect for the upper limit of such a quantity. Still, it provides a very useful quantitative scale describing one of the most important properties of plating baths. In the case of purely primary distribution, where $Wa \ll 1$ and there is effectively no throwing power, the thickness of the deposit will simply be inversely proportional to the distance. This yields $M = K$ and $\mathrm{TP} = 0$, as expected.

Having defined the throwing power quantitatively, we can now proceed to discuss the physical reasons for the dependence of the throwing power on geometry and the methods available to increase the value of the TP to acceptable levels.

18.1.7.2 Micro Throwing Power

In the previous section, we discussed *macro* throwing power, which determines the uniformity of plating. Here, we shall discuss the *micro* throwing power, which controls the smoothness and brightness of the deposit, and naturally depends on quite different factors.

Scale is very important in electrode processes. In the case of *macro* throwing power, the irregularities of the shape of the electrode are on the same scale as

the cell itself. The distance between the anode and the cathode may be of the order of 1–10 cm, and the characteristic length used to calculate the Wagner numbers would be in that range. When the appearance of the surface, particularly its brightness, is considered, the scale of interest is of the order of magnitude of the wavelength of visible light. It follows from electromagnetic theory that the ratio between the light scattered from a surface and that reflected from it depends on the ratio $(h/\lambda)^2$, where h is the amplitude of the roughness and λ is the wavelength of light. If $(h/\lambda)^2$ approaches zero, one has specular reflection; that is, the surface reflects light like a mirror. As this ratio grows, the surface first looks dull, and eventually becomes black as $(h/\lambda)^2$ exceeds unity. Thus, the scale of interest, from the point of view of brightness, is of the order of magnitude of the wavelength of visible light, namely, in the range of 0.4–0.8 μm. The ability of a plating bath to form uniform coatings on this scale of roughness is called *micro throwing power*.

We can see intuitively that cell geometry has little to do with micro throwing power. If the amplitude of roughness is of the order of 1 μm or less and the distance between the anode and the cathode is a few centimeters, the variation of solution resistance at crests and valleys on the surface must be negligible. Hence, primary current distribution should not have any effect on the micro throwing power. A good way to look at it is to consider the appropriate Wagner number for this situation. Setting the characteristic length in Eq. (18.4) at a few micrometers, instead of a few centimeters, the value of Wa will increase by a factor of about 10^4, which is deep in the range of secondary current distribution.

There must be a different mechanism controlling micro and macro throwing power. This is tertiary current distribution, which is mass-transport-limited. A common observation in electroplating is that the roughness of the deposit increases with thickness. It is quite easy to produce a smooth deposit of 0.1 μm thickness, but keeping it smooth when the thickness has grown to 100 μm requires very special measures. Such an observation implies that a positive feedback mechanism is in operation, with the local current density higher at protrusion than in recessed areas. It is easy to understand this behavior, if the plating process is assumed to be at least partially controlled by mass transport. We recall that the current density can be written as

$$j = nFD(c_b - c_s)/\delta \tag{1.3}$$

In industrial operations, stirring is typically implemented by moving the electrodes, stirring the solution, or pumping it through the bath (the latter provides an opportunity to filter the solution and remove particulate material that could damage the quality of the metal coating). The typical values of δ, under industrial plating conditions, may be 50–150 μm, placing it between the values relevant for primary and secondary current distributions.

The Nernst diffusion layer thickness is larger in a recessed area than at a crest; hence, the local current density is smaller. As a result, recessed areas grow more slowly than crests, and the amplitude of roughness increases with time and, hence, with thickness during plating by a positive feedback mechanism.

It is not difficult to see how a rough surface will grow even rougher by the foregoing mechanism, but how is roughness initiated? Experiments show that

even when plating is conducted on a highly polished surface, the deposit will gradually increase in roughness. We may expect that plating on an atomically flat, single-crystal surface in a highly purified solution will not produce a rough deposit, but this is of little practical interest. In practice, a surface is prepared by degreasing, activation by acid (to remove oxides), and so on, but rarely by polishing. Indeed, the surface is often roughened in order to improve adhesion, but even a polished surface will not be completely flat, so that there will always be irregularities that can serve as preferred sites for initiating the above positive feedback mechanism. In addition, real plating baths usually contain some foreign particles that adhere to the surface during plating. These contaminants may be dust particles or solid grains of metal that fell from the anode during plating. It should be remembered here that impurity particles of sub-micrometer dimensions, which are often difficult to remove by filtering, are fairly large on the scale of importance here.

Another mechanism of roughness initiation may be associated with the nonuniformity of the substrate. The activation-controlled current density at sites of inclusions (such as graphite or sulfur), at grain boundaries, and at different crystal faces may be different, causing uneven growth of the deposit.

18.1.8 The Use of Additives

There is great commercial incentive to produce smooth and bright deposits. Consequently, there is a vast choice of additives for use to improve *micro* throwing power, making deposits smoother and more uniform and producing a bright metal luster. Although the properties of additives differ widely, the mechanism by which they operate is common and easy to understand.

Molecules of the additive adsorbed on the surface prevent or inhibit metal deposition. To a first approximation, it can be said that the rate of metal deposition is simply proportional to the fraction of the surface that is not covered by the additive. A more detailed analysis shows that adsorption on part of the surface could also have an effect on the rate of metal deposition on the bare sites, but this refinement need not concern us here. As a rule, the concentration of the additive is very small compared to that of the metal ion being plated. Consequently, the rate of adsorption of the additive is controlled by mass-transport limitation, while the rate of metal deposition is mostly activation controlled, with possibly some mass-transport limitation involved, depending on the ratio of j/j_L, where j is the fractional current density consumed for deposition of the metal. This helps to produce a smooth surface for the same reason that a rough surface is formed in the absence of a suitable additive. On protruding parts on the surface, the rate of mass transport is higher than on flat or recessed regions. The additive is preferentially adsorbed on such regions, preventing, or at least inhibiting, metal deposition. As a result, the current density for metal deposition is higher in the recessed or flat regions, leading to leveling. Under favorable conditions, this effect can actually be strong enough to reverse the trend, namely, to yield a smooth deposit on an initially rough surface.

Some of the additive adsorbed on the surface may be buried as such. Alternatively, an additive may first be reduced, whereupon fragments of it are buried

under the layers of metal being deposited. In either case, the additive is consumed during operation of a plating bath and must be periodically replenished. The incorporation of foreign molecules in the metal deposit affects its mechanical properties, as well as its corrosion resistance. These effects cannot generally be predicted by theory, and here the art of finding the right additive for each plating bath comes into play.

There is an optimum range of concentrations over which each additive is most active. This is also easy to understand, in terms of the mechanism just discussed. At low concentrations, the activity of each additive grows with increasing concentration, because there is not enough material in solution to do the job; that is, coverage on the protruding areas cannot reach a sufficiently high value to induce significant leveling. Within the best concentrations range, coverage on protruding areas is high, but in recessed areas it is relatively low, yielding the desired leveling effect. As the concentration of the additive in solution is increased further, the coverage on protruding areas reaches a limiting value and can grow no longer; the coverage on other areas keeps growing, however, until a high coverage is reached everywhere on the surface.

The effect of stirring on the rate of adsorption of an additive can be utilized to identify the optimum range of concentrations for any particular system. This can best be performed with the use of a rotating disc electrode (RDE), for which the rate of rotation can be scanned while the current density is maintained constant. If the concentration of the additive is optimal, the measured overpotential will increase with increasing rate of rotation, because the rate of supply of additive to the surface is increased, leading to stronger inhibition of metal deposition. If the concentration of the additive is too high, surface saturation is already reached at low rotation rates, and increasing the rate of mass transport may have the effect of decreasing the overpotential for metal reduction, if this reaction is partially mass-transport-controlled.

We may ask what determines the suitable range of concentrations of an additive. It is clear that the answer is different for different additives and depends on the metal being deposited. If the adsorption isotherm for the additive on the same metal is known, a good guess would be to use a concentration of the additive that will lead to a partial coverage, in the range of $0.2 \leq \theta \leq 0.8$, because this is the range in which $\partial\theta/\partial c$ is the greatest, and one may expect to obtain the highest difference in adsorption on different regions on the surface. On the other hand, if the adsorption isotherm is not known, which is usually the case, it is probably easier to determine the optimum concentration range experimentally by trial and error than to measure the isotherm and deduce the desired range of concentrations from it.

In a large majority of practical plating baths, a *complexing agent* is used to improve the quality of the product, in particular to obtain smooth and bright deposits. On the other hand, when deposition at a high rate is needed, as for electroforming, the metal is deposited from a simple solution containing no complexing agent. Until about the middle of the twentieth century, the most commonly used complexing agent was cyanide. Many metals such as Au, Ag, Cu, Ni, Co, Cd, and Zn were plated from alkaline baths containing KCN. This practice was abandoned for environmental considerations, in spite of the fact that it was most

satisfactory from the purely engineering point of view. Cyanide has been replaced by other complexing agents, mostly organic polyacids, having two or more carboxylic groups, such as citrate and oxalate. The main purpose of formation of the complex is to slow down the kinetics of the electrodeposition reaction. It should be noted that in most industrial electrolytic processes, such as the production of metals, the chlor-alkali industry, and in organic electrosynthesis, as well as in batteries and fuel cells, fast kinetics is an advantage – decreasing the overpotential and thereby reducing energy consumption. The opposite is true in metal deposition. Decreasing the exchange current density is equivalent to increasing the faradaic resistance. This generally leads to improved uniformity of the deposited layer and enhanced smoothness and brightness of the deposit.

This raises some important possibilities, which have not escaped the attention of the electroplating community. For example, while metal deposition is conducted in fairly concentrated solutions of the metal being plated, and at current densities well below the mass-transport limit, additives acting as inhibitors for metal deposition are often introduced at concentrations that are several orders of magnitude lower, to ensure that their supply to the surface will be mass-transport-limited. In this way, the tendency for increased rate of metal deposition on certain features on the surface, such as protrusions, will be moderated by the faster diffusion of the inhibitor to the very same areas. Furthermore, if deposition occurs in the region of mixed control, which is usually the case, it must be remembered that the relevant roughness factor is quite different for the charge-transfer and the mass-transport processes, and this may well be a function of current density, since the faradaic resistance is inherently potential dependent.

The choice of a good leveling agent depends, among other things, on the excess charge on the metal, q_M, at the potential where metal deposition takes place. The latter depends, of course, on the potential of zero charge of the metal.[2] A positively charged additive will be preferentially adsorbed on a negatively charged surface, and vice versa. A neutral additive will be adsorbed mainly around the point of zero charge (see Chapter 12). The potential at which the metal is actually deposited depends on its standard potential, its concentration, and the complexes it forms with ligands in the solution, as well as on the rate of deposition.

Although the mechanism by which different additives operate is fairly well understood, we have certainly not reached the point at which the choice of an additive can be based on its known molecular structure, or even on measurement of its adsorption isotherm under equilibrium conditions. Such knowledge can be used to advantage for preliminary screening and intelligent guessing, but it cannot eliminate some degree of trial and error in identifying a good additive for a given purpose.

Before concluding this section, a short discussion of the *Hull cell* (R.O. Hull, *Am. Electroplat. Soc.* **27** (1939) 52), which is commonly used in industry, should

2 We need to be careful here in deciding "which metal." At the very beginning of plating on a foreign substrate, it will be the value of E_{pzc} for the substrate, but when the thickness of the deposit exceeds several nanometers, it will be the value E_{pzc} for the metal being plated, because there is no longer any contact between the solution and the substrate.

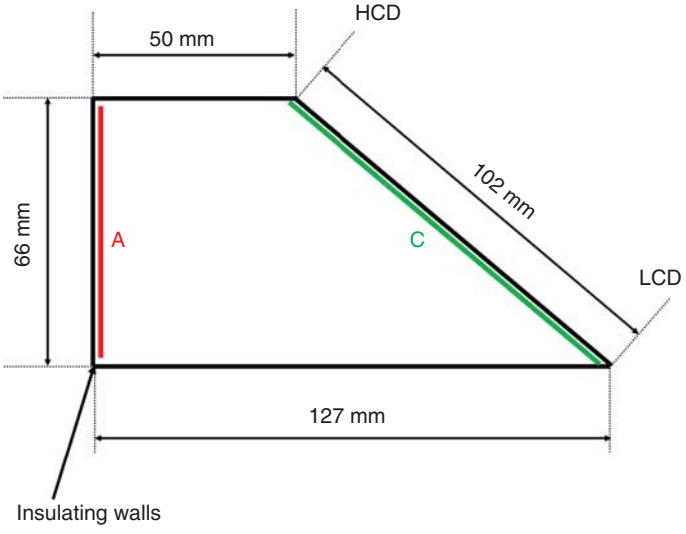


Figure 18.4 Schematics of the 267-mL Hull cell.

be provided. This is a trapezoidal plating test cell, generally made of an insulating material such as Perspex, polypropylene, or PVC, in which the cathode is positioned at an oblique with respect to the anode (see Figure 18.4). Consequently, a wide range of current densities on the cathode can be studied in a single experiment, the highest current density being at the point closest to the anode (HCD).

A useful empiric approximation of the current distribution for a 267-mL cell is

$$j(x)/I = 51.0 - 52.4 \log(x) \tag{18.15}$$

where I is the total current (A), x is the distance along the cathode from the HCD edge (cm), and $j(x)$ is the local current density (mA cm^{-2}) at distance x. It should be noted that this formula does not take into account the influence of electrochemical kinetics.

The Hull cell is usually used with a magnetic stir bar on the bottom, which results in stronger agitation at the bottom of the cathode panel than at the top. After plating, the panel is removed and visually inspected for the effects of variations in current density and agitation on the deposit morphology or brightness. The Hull cell is also commonly used to analyze the effects of additive concentration and to check the condition of the electroplating bath. It allows determining proper operating windows. Moreover, by making all additions to the Hull cell before adding anything to the main bath, one can avoid making mistakes that might result in irreparable damage to the bulk of the electrolyte.

18.1.9 The Microstructure of Electrodeposits and the Evolution of Intrinsic Stresses

Electrodeposition, by definition, is a nonequilibrium process as it occurs under a cathodic overpotential. Consequently, the phases formed are often

not in thermodynamic equilibrium, and may be either undersaturated or supersaturated compared to their counterparts in the equilibrium phase diagram.

The structures of electrodeposits are mainly classified as (i) fine-grained (mostly equiaxed); (ii) columnar; (iii) fibrous; or (iv) banded. Columnar structures are characteristic of deposits from additive-free solutions (especially acidic), which are formed at a low current density (i.e. low deposition rate) or elevated temperatures. They are typically characterized by lower tensile strengths and microhardness values than other structures, but are generally more ductile and have high purity and density, and a small electrical resistivity. In contrast, banded structures are characteristic of brightener-containing solutions, and processes that involve either multiple baths or periodic variations of some deposition parameter (e.g. pulse plating). Spontaneous formation of banded structures has, however, also been observed in the electrodeposition of some pure metals (e.g. Ni) and some alloys (e.g. Ni–P and Ni–Co). The layers often have a submicron thickness, and are thought to be compositional or structural modulations that originate from small-scale periodic fluctuations in the metal or additive concentrations, pH, agitation, or polarization at the cathode's surface. Banded structures generally possess higher tensile strengths and microhardness values, but also higher internal stresses and lower ductility than other structures. Additionally, compositionally modulated multilayers have been reported to exhibit remarkable wear resistance, magnetic, electronic, superconducting, and optical properties.

The initial electrodeposited layer inherits the structure or orientation of the substrate when the misfit strain energy between this layer and the substrate is less than 12%. After this epitaxial layer achieves a critical thickness (1 μm or less), fine columnar grains extending along the growth direction may form, and the deposited crystallographic texture then becomes independent of the substrate structure. These microstructures are strongly dependent on the different deposition conditions, such as the plating bath composition and temperature, applied current density, and orientation and surface morphology of the substrate. Several theories have been proposed to explain the development of microstructures in electrodeposits. The microstructure often appears to be the outcome of a combination of different mechanisms.

The effects of the adsorption of hydrogen atoms on the texture of the coating have been suggested for columnar and lamellar growth during electrodeposition of Ni. Free growth can be related to a substrate surface in the absence of adsorbed hydrogen atoms and surface-active substances. For single crystals, different lattice planes have different free growth metallic overpotentials; the higher the reticular density (i.e. number of atoms per unit area on the face), the lower is the growth rate in a direction normal to the face. Hydrogen atom intermediates are adsorbed more easily on lattice planes, which during free growth have faster growth velocities. Thus, hydrogen adsorption results in a greater free energy of activation for Ni deposition, i.e. an increase in the metallic overpotential above the free growth value. The texture is affected by the average surface coverage of hydrogen atoms (θ_{av}); the higher the value of θ_{av}, the higher is the hydrogen content of the electrodeposit. Different values of θ_{av} result in different textures.

The most important parameters determining the mode of growth of a substance on a substrate made of another material are the deposit/substrate

binding energy and the lattice parameter and structural misfit between them. For a small supersaturation and negligible kinetic influences, a deposit may grow on a substrate via different modes. The Volmer–Weber (VW) mode is preferred, regardless of the lattice parameter and structural misfit, when the binding energy between adsorbed deposited ions is greater than the binding energy between an adsorbed ion and a substrate atom. The formation of nodules (colonies) may be the outcome of nucleation extending laterally from an existing structural defect (e.g. dislocation) and sweeping impurity atoms (e.g. hydrogen) ahead of the growth front. The impurity atoms would therefore be more concentrated where nodules meet and inhibit outgrowth, i.e. perpendicular to the surface. The mesoscale colony surface structure is also known as a nodular or cauliflower structure, and is often related to a VW growth mode.

Surface crevices (or cracking) in electrodeposits may be related to decomposition of unstable hydrides during or soon after plating, which involve significant volumetric contractions and internal stresses. Examples include decomposition of hexagonal nickel hydride Ni_2H to pure hexagonal or face-centered cubic (fcc) Ni, or decomposition of hexagonal chromium hydride to body-centered cubic (bcc) chromium. Even in the absence of hydrides, codeposition of hydrogen may result in intrinsic stresses, either tensile (if hydrogen diffuses from the surface layer to the layers beneath it) or compressive (if hydrogen remains in the surface layer and fills voids). Such macroscopic intrinsic stresses are likely to be larger in plating systems with a low FE. Residual stresses may also be related to the manner in which metal grains crystallize (e.g. variation in grain size), or to defects incorporated in the coating (e.g. microscopic voids, metallic impurities, dislocations, and interfaces). Macroscopic extrinsic stresses are mainly due to epitaxial growth and a significant difference between the interatomic spacing of the substrate and deposit materials. When the deposit thickness reaches a critical value, misfit dislocations start developing, thereby relieving these stresses. Macroscopic intrinsic stresses develop within the deposit when it is restrained from shrinking or expanding. For example, when 3D epitaxial crystallites coalesce before the spaces between them have been completely filled, the surface layer of a deposit tries to shrink. The restraint by the layers under the surface causes tensile stresses, which become larger as the grains become smaller.

Abermann (*Vacuum*, **41** (1990) 1279) classified the time (thickness)-dependent stress buildup patterns in VW vapor-deposited thin films into two types. Type I is typical of high melting point elements, e.g. Re, W, and Ta. For these elements, the final microstructure often has high-aspect-ratio columnar grain structure. Tensile stresses are developed during growth, but when growth stops, the final film's stress state usually does not change, or changes only slightly, with increasing time. The stress buildup in electrodeposited films is, however, significantly different from that occurring in vapor-deposited films.

18.1.10 Pulse Plating

In the past three decades or so, specific interest in *pulse plating* and *periodic reverse pulse plating* (where the current is either interrupted or reversed periodically, respectively) has increased significantly. These techniques may have a significant influence on the composition and structure of electrodeposited alloys as well as on the surface morphology and the micro throwing power.

Application of a pulsed current can allow for current densities, faradaic efficiencies, compositions, and microstructures that are unattainable with DC plating. Pulse plating has widely been used to increase the uniformity, adhesion, and speed of electrodeposition as well as to generate unique morphologies. The high overpotentials made possible by pulse electrodeposition increase the available energy for nuclei formation, and therefore higher nucleation rates are attainable with pulses, as compared to DC plating. Pulse electrodeposition can therefore be used to produce coatings with finer grains.

Pulse plating plays a critical role in the mass transport of ions that exist in dilute concentrations. Application of a pulsed current modifies the transient concentration gradient profile of the diffusion layer by allowing for regeneration of depleted ions from the bulk during the relaxation time. This creates a steeper concentration gradient and allows for higher partial current densities within a certain range of pulse parameters. If, however, the concentration of the metal ion is large enough, its concentration profile will be insensitive to the pulse parameters and will always resemble the profile in the bulk.

Among the important parameters of pulse plating are the total pulse time, $\tau_{pp} = t_{on} + t_{off}$, and duty cycle, $\theta = t_{on}/(t_{on} + t_{off})$. Through replenishment of the diffusion layer during t_{off}, pulse electrodeposition allows for a higher instantaneous current density by creating a steeper concentration gradient. During t_{off}, electroless adsorption, desorption, and recrystallization reactions can occur on the surface as a result of relaxation at the interface.

In a pulsing system, the average current density, j_{avg}, of a species can at most equal its DC limiting current density, j_L:

$$\frac{j_p t_{on}}{\tau_{pp}} = j_p \theta = j_{avg} \leq j_L \tag{18.16}$$

where j_p is the applied peak current density. If j_{avg} exceeds j_L, side reactions (e.g. hydrogen evolution) will necessarily be observed and the FE must be below 100%. The peak current density, j_p, that satisfies the condition $j_{avg} = j_L$ is the steady-state limiting peak current density of a pulsing system. From Eq. (18.16) it follows that an increase in either the on-time or the peak current density leads to an increase in j_{avg}. It also follows from Eq. (18.16) that as the duty cycle decreases, the steady-state limiting current density increases. This limit, however, is rarely actualized as non-steady-state mass-transport limitations usually set in before the steady-state limit is reached.

In systems where pulsing gives rise to changes in the diffusion layer profile of the electroactive species, the limiting current density is instead characterized by the pulse limiting current density, j_{pL}, which takes into account the transient effects of the pulsating diffusion layer not described by the j_L variable. The pulse limiting current density represents the current density for which the surface concentration of electroactive species equals zero at the end of a pulse. It is generally described by its non-steady-state limitation due to the pulsing diffusion layer adjacent to the electrode surface. In most systems, only at short pulse times does j_{pL} approach the steady-state limit.

The j_{pL} is among the most important quantities used to describe a pulsing system. It represents the current value required to bring the surface concentration of a species to zero by the end of the pulse. It follows then that shorter pulses have larger values of j_{pL} at a constant duty cycle. Several models are present in the literature to describe the effect of pulse parameters on j_{pL} for single-component systems. The pulse limiting current density, j_{pL}, of a single component under diffusion control is given by

$$j_{pL} = j_L \frac{1}{1 - 2t^* \sum_{m=1}^{\infty} \frac{\exp[\lambda_m(1-\theta)]-1}{\lambda_m[\exp(\lambda_m)-1]}} \tag{18.17}$$

$$t^* = D\tau_{pp}/\delta^2 \tag{18.18}$$

where t^* is the dimensionless pulse time, and δ is the boundary layer thickness in centimeters, and λ is a dimensionless summation parameter given by

$$\lambda_m = \pi^2 t^* \left(m - \frac{1}{2}\right)^2 \tag{18.19}$$

18.1.11 Plating from Nonaqueous Solutions

18.1.11.1 Statement of the Problem

Many metals can be plated from aqueous solutions, even though their reversible potential is cathodic with respect to the RHE in the same solution. Hydrogen evolution can occur in such cases as a side reaction, but as long as the current efficiency is not too low, plating can be conducted on an industrial scale. One of the important reasons for this is that the exchange current density for metal deposition is usually much higher than that for hydrogen evolution, with the result that the rates of these reactions are comparable, even if the reversible potential for metal deposition is significantly more cathodic. In the presence of a suitable ligand, forming a complex with the metal ion being electroplated, the rate of metal deposition is slowed down (to achieve better throwing power) and the potential of deposition is shifted cathodically, but the high concentration of the ligand at the metal surface during plating can lower the rate of hydrogen evolution, allowing the process to occur at a reasonable current efficiency. Also, such electroplating baths usually operate in neutral or somewhat alkaline pH, and the reversible potential for hydrogen evolution is shifted in the negative direction, decreasing the overpotential for hydrogen evolution at the potential where plating is conducted.

More active metals, such as Al, Ti, and Mg, cannot be deposited from an aqueous medium at all. An attempt to do so leads to copious hydrogen evolution, but no detectable metal deposition. On the other hand, Na and other alkali metals can be deposited on Hg from an alkaline solution, probably because of the very low exchange current density for hydrogen evolution on this metal, and because an amalgam is formed, so that the active surface is always Hg or its amalgam, not the metal being deposited.

When electroplating from an aqueous solvent is impossible, one must resort to nonaqueous systems. These present a number of technical difficulties and have

been used in practice only when there has been no alternative. With evolving technological development, it is anticipated that plating from nonaqueous systems may nevertheless be adopted for commercial use, and a short discussion is therefore warranted.

Since hydrogen evolution via decomposition of water is the problem, the most obvious way to proceed is to plate from a nonaqueous solution. Moreover, the solvent used should not have an active proton that could readily allow hydrogen evolution, such as an alcohol. A molten salt, such as $MgCl_2$, would seem a good solution to the problem. Indeed, metallic Mg is synthesized from anhydrous molten $MgCl_2$, and similarly Al is produced from a bath containing bauxite (Al_2O_3) dissolved in molten cryolite (Na_3AlF_6) at high temperatures, since in these baths metal deposition is the only cathodic reaction that can take place. The quality of the deposits is usually poor, however, and the high temperature of operation (particularly in the case of Al) could damage the substrate. Hence, high-temperature molten salts are suitable for synthesizing the relevant metals, but rarely used for plating them as a thin-layer coating.

Refractory metals, such as Ta and Zr, can be deposited from their fluorides in a molten salt bath. In the case of Zr, for example, the bath consists of ZrF_4 or ZrF_6^{2-} in a KF/NaF/LiF mixture. The alkali fluorides are employed to increase conductivity and decrease the melting point. Even so, these baths are operated at about 800 °C. Good deposits have been reported as long as the right valence was chosen for each metal (3 for Mo and V, 4 for Nb and Zr, 5 for Ta). The bath must be operated in a pure argon atmosphere, and impurities must be strictly excluded. It should be obvious that the operation of such baths is expensive, and their control is difficult. Thus, their use is limited either for research purposes or for highly specialized applications, where cost is of secondary concern.

Molten-salt electrodeposition allows obtaining refractory compounds such as borides and carbides, which cannot be electrochemically deposited from aqueous solutions. In addition, it enables the formation of the metallurgical bonding between the coating and the substrate due to the high-temperature diffusion process; thus, high adhesion strength can be achieved.

The electrodeposition of semiconductors has attracted much interest in recent years, primarily because this is a relatively simple and inexpensive deposition technology that may be scaled up easily. Although the electrodeposited films do not exhibit the crystalline perfection or low levels of electrically active impurities as single-crystal epitaxial films deposited by molecular beam epitaxy, CVD, and some other techniques, there are applications such as photovoltaic power generation where large areas of semiconductors are required and the properties of the electrodeposited films are good enough. In addition, for application in solar cells electrodeposition enables to easily alter both the bandgap and lattice constant by composition modulation. The electrodeposition of Si must employ a nonaqueous medium because of its very high reduction potential and the high reactivity of most of its compounds to water. Silicon films have been deposited, among others, from LiF–KF–NaF–K_2SiF_6 melts. GaAs has been deposited from molten salts such as B_2O_3–NaF–Ga_2O_3–$NaAsO_2$.

18.1.11.2 Methods of Plating Al

The search for a room-temperature plating bath for aluminum has been conducted for many years, in view of the excellent corrosion resistance and low toxicity of this metal. An early technological success was the so-called hydride bath, which consisted of a solution of $AlCl_3$ and $LiAlH_4$ in di-ethyl ether, $(C_2H_5)_2O$. A large excess of $AlCl_3$ was used ($AlCl_3$:$LiAlH_4 = 7:1$). The exact mechanism of Al deposition from this bath is not known, but there can be no doubt that the negative hydride species in $LiAlH_4$ plays a crucial role, since the bath cannot be operated after it has been depleted of $LiAlH_4$, even if the concentration of $AlCl_3$ is kept constant. This technology has been used on a few occasions for highly specialized purposes, mainly for the production of aluminum mirrors that are excellent reflectors of infrared radiation. It has not gained widespread commercial application because it requires the use of a highly flammable solvent and chemicals that are very sensitive to water and oxygen.

Aluminum can also be plated from a low-temperature molten bath employing a mixture of $AlCl_3$ and KCl. The melting point depends on composition, and the bath can be operated in the range of 200–300 °C. Two anions can exist in this melt: $AlCl_4^-$ and $Al_2Cl_7^-$. Their relative concentrations depend on the ratio of the two salts used. The great advantage of this system is that it dissolves chlorides of other metals, such as Ti and Mn, and allows the deposition of alloys of these metals with Al. The greatest disadvantage is that the bath must be operated under strictly anhydrous conditions, since $AlCl_3$ is highly hygroscopic, releasing HCl when in contact with water or humid air.

Another near-room-temperature bath for Al plating contains a metal-organic compound, $Al(C_2H_5)_3$, dissolved in toluene, with $AlCl_3$ added to provide electrolytic conductivity. This bath, operated at about 1000 °C, has excellent conductivity and good throwing power. Although it has made some inroads to engineering applications, its widespread application has been limited, probably because of the need to use an expensive and dangerous metal-organic compound that ignites spontaneously in air.

A room-temperature plating bath is based on the use of Al_2Br_6 and KBr in toluene, ethyl-benzene, or similar aromatic solvents. The chlorides and bromides of Al are covalent compounds and are highly soluble in aromatic hydrocarbons. An ionic compound such as KBr is not soluble in an aromatic hydrocarbon, but is readily dissolved in a solution of Al_2Br_6 in toluene, forming a complex ion

$$Al_2Br_6 + KBr \rightarrow [K^+(Al_2Br_7)^-] \tag{18.20}$$

Ionization of the species formed in solution would still lead to the formation of potassium ions, which are unstable in aromatic solvents. The real ionic species in solution must therefore be somewhat more complex. Experiments indicated that the actual species in solution are

$$3[K(Al_2Br_7{}^-)] \rightleftharpoons [K_2(Al_2Br_7)]^+ + [K(Al_2Br_7)_2]^- \tag{18.21}$$

This bath has the advantage of operating at room temperature, but it is also very sensitive to moist air, forming highly corrosive and toxic HBr, and it has not been adopted by industry.

It is interesting to note that all three aluminum plating baths discussed here employ a solvent of low polarity. In fact, Al cannot be deposited from any polar solvent. For example, a solution of LiCl and $AlCl_3$ in acetonitrile or propylene carbonate yields a deposit of metallic Li, but no Al, even though thermodynamically Al should be deposited first. The reason evidently lies in the kinetics of the process. In any polar solvent the energy of solvation of the small Al^{3+} ion is so high that the first step in the reaction sequence – the removal of a single solvent molecule from the inner solvation shell – requires a very high energy of activation. It is only when the solvent is nonpolar that this process can proceed at a significant rate.

It is not easy to find a suitable nonaqueous electrolyte for plating Al, Ti, and other active metals. Operating such a bath may be even more difficult. First, water and often oxygen must be excluded. This can be done rather easily in continuous processes, such as plating a wire or a metal sheet. In most applications, however, electroplating is typically a batch process, i.e. parts are introduced and removed from the bath routinely. While the technology to perform such operations exists, it is more expensive and much less convenient than operation in an open aqueous bath. Most nonaqueous solvents are either flammable or toxic, or both. Most salts used to make up the bath are expensive, and some are quite unstable. Even a relatively inexpensive salt such as KBr can become expensive when it must be kept dry. Moreover, a current efficiency that is just a little below 100%, which may be a minor irritation in aqueous solution, could turn out to be a major problem in nonaqueous media, because the side reactions can lead to the accumulation of products that are detrimental to the operation of the bath, to say nothing of the health of the operator. Waste disposal, a problem even in aqueous plating baths, can be much more problematic and costly in nonaqueous baths.

These are some of the reasons for the failure of nonaqueous plating baths to come into general use, in spite of some clear technological advantages they can offer, in particular in making products that cannot be made by other means. There is little doubt, however, that rigorous specifications of emerging new technologies, accompanied by research and development in this field, could eventually lead to the introduction of nonaqueous plating technologies into industrial applications.

18.1.12 Electroplating of Alloys

18.1.12.1 General Observations

There is much in common between the electroplating of a single metal and of an alloy of two or more metals. Both are inner-sphere charge-transport processes, characterized by the fact that the initial state is a strongly hydrated ion in solution and the final product is a bare atom that has become a part of the lattice of the metal electrode. In both cases, charge is carried across the interphase by movement of the metal ion(s), and in both cases the surface of the substrate changes as the process goes on. In both cases, hydrogen evolution is a common side reaction, so that the current efficiency could be less that 100%, and moreover, it could be a function of the current density applied. However, there are also significant differences that have to be taken into account when studying alloy deposition:

(1) Alloy deposition, by definition, must involve at least two simultaneous electrochemical reactions occurring in parallel. In the general derivation of the equations for electrode kinetics one obtains an expression of the form

$$j = nFkc_b \left[\exp\left(\frac{\alpha_a F\eta}{RT}\right) - \exp\left(-\frac{\alpha_c F\eta}{RT}\right)\right] \tag{18.22}$$

where it is tacitly assumed that only a single reaction takes place and the composition and morphology of the surface are invariable during the experiment.[3] If we try to apply this equation to deposition of a binary alloy, three problems immediately emerge. First, the concentration of the two metals in solution is not necessarily the same. Indeed, in most cases they have to be different in order to achieve the desired composition of the alloy being deposited. Secondly, at any given applied (or measured) potential, the overpotentials are different for the two metals. Thirdly, the transfer coefficients (and, consequently, the cathodic Tafel slopes of the two metals) can be different. Consequently, measurement of the current–potential relationship does not allow us to draw any simple interpretation of the mechanism involved, although it does show in some cases the potential where deposition of each of the two metals starts. Even if the experimental data can be "forced" to provide a linear Tafel region, from which an apparent Tafel slope is obtained, this will have little relevance to the mechanism of formation of the alloy. Although the above observation is well known and could be considered rather obvious, it has been overlooked in many publications.

(2) Equation (18.22) is valid only if the reaction occurs under purely activation control. When a single metal is being deposited, mass-transport limitation can be taken into account by replacing the bulk concentration in Eq. (18.22) with the surface concentration. At steady state this is rather simple, since

$$c_s/c_b = 1 - j/j_L \tag{4.14}$$

During alloy deposition the concentrations of the two or more alloying elements in solution could be different – one may be deposited at a rate controlled substantially by mass transport (and, hence, not influenced much by the applied potential), while the other may be essentially under conditions of activation control. We should also take into account the number of electrons needed to reduce each of the alloying elements to the neutral metal atom. For example, in deposition of a Au–Ni alloy from a solution containing $AuCl_4^-$ and $NiCl_2$, at equal concentrations, the limiting current for Au would be 50% higher than that for Ni (if the difference in the diffusion coefficients of the two ions is ignored).

(3) An attempt to calculate the extent of mass-transport limitation for each of the alloying elements should be based on the partial current density for that element, and never on the total current density applied. Unfortunately, this quantity cannot be derived from the current–potential relationship itself. It has to be calculated from the FE (which is the measure of the rate of hydrogen evolution) and an analysis of the atomic composition of the alloying elements

3 Bear in mind that $\alpha_a + \alpha_c = n$; the two transfer coefficients do not generally add up to unity.

in the deposit. Thus, the best we can do is to determine the average value of the partial current density of each alloying element. Admittedly, this limitation also applies to deposition of a single metal, unless one can be sure that hydrogen evolution does not occur at the potential where the deposition process takes place.

(4) The partial currents for deposition of each of the alloying elements should be determined as a function of potential and other experimental parameters via determination of the atomic composition of the alloy and the FE. The FE during alloy deposition can be different from that of single-metal deposition of one or both metals involved in the process. Hence, the FE can be expected to depend on the composition of the alloy, and the thickness distribution may differ from that expected according to the current distribution.
The FE can be calculated using the equation

$$\mathrm{FE} = \frac{w}{It_\mathrm{d}} \sum \frac{c_i n_i \mathrm{F}}{M_i} \times 100 \tag{18.23}$$

where w is the measured added mass of the deposit (g), I is the total current passed (A), t_d is the deposition time (seconds), c_i is the weight fraction of the element in the alloy deposit (determined by SEM–EDS, for example), n_i is the number of electrons transferred per atom for the reduction of the specific element, F is Faraday's constant, and M_i is the atomic mass of each metal ($\mathrm{g\,mol^{-1}}$).
The partial current density ($\mathrm{mA\,cm^{-2}}$) for element i is then defined by

$$j_i = \frac{w}{At_\mathrm{d}} \frac{c_i n_i F}{M_i} \tag{18.24}$$

where A is the surface area of the cathode ($\mathrm{cm^2}$).

(5) Uniformity of current distribution is always a problem in electroplating, but in alloy plating there could be an added complication – in addition to leading to nonuniform thickness, it could also lead to nonuniform composition, since the rate of deposition of alloying elements may depend differently on the local current density.

(6) Changing the concentration of one of the metals in solution can bring the reversible potentials closer to each other. However, since the Nernst equation allows for only $(59/n)$ millivolts per decade change in concentration, where n is the number of electrons needed to deposit a metal atom, this has a limited range of applicability. When the two metals forming the alloy have widely different values of E^0, their reversible potentials can be shifted closer to each other by adding a complexing agent that forms complexes with different stability constants.

(7) Deposition of alloys is usually conducted in the presence of one or more ligands. For example, in plating of a Ni–W alloy, both citrate and ammonia are added to the solution. Now, citrate forms two different complexes with Ni, and NH_3 could form as many as five different complexes of the type $[Ni(NH_3)_n]$. In addition, the WO_4^{2-} and the Ni^{2+} ions can form two types of complexes with citrate. Adding the hydrogen evolution reaction, this could

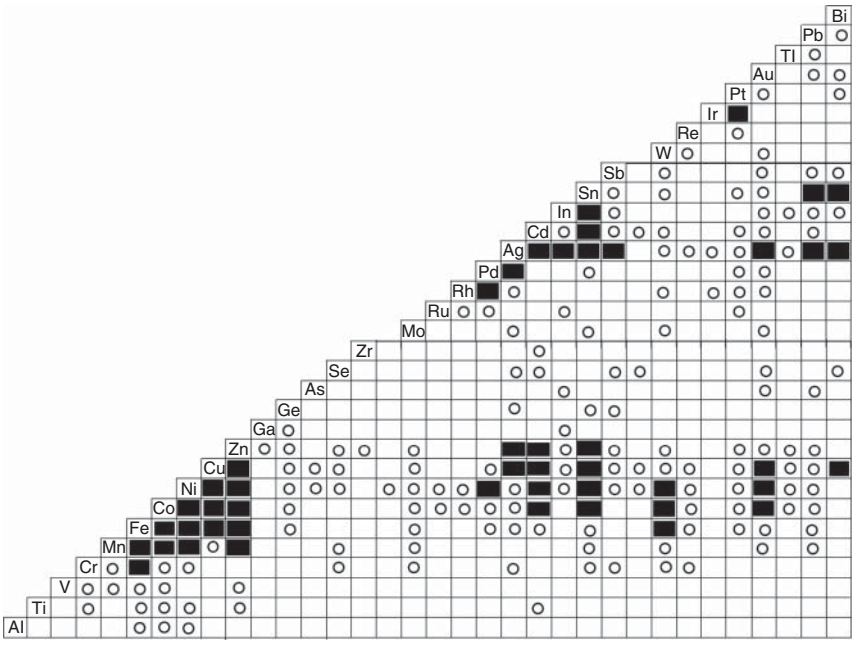


Figure 18.5 Metal pairs that have been codeposited electrochemically: ○ = Demonstrated in laboratory studies and/or small-scale applications, ■ = Technically interesting, widely employed alloys. Source: Reprinted with permission from N. Kanani, *Electroplating: Basic Principles, Processes and Practice*, Elsevier, Oxford, 2004.

add up to as many as 10 cathodic reactions taking place in parallel simultaneously!

18.1.12.2 Some Specific Examples

Figure 18.5 shows many of the metal pairs that have been codeposited electrochemically, either commercially or in laboratory studies.

Simple alloy deposition is observed for the Pb–Sn couple. The standard potentials of the two metals are −0.126 and −0.136 V versus SHE, respectively; hence, it is easy to adjust the concentrations in solution so that the reversible potential would be the same, or close to each other. Although plating occurs at a potential negative with respect to the reversible potential for hydrogen evolution, both metals are very poor catalysts for this reaction, so that close to 100% FE is readily obtained. The most important consequence of this is that the composition of the plated alloy is linearly related to the composition of the solution. This is, unfortunately, not common in alloy plating, as we show next.

Anomalous codeposition (ACD) describes an electrochemical deposition process in which the less noble metal is deposited preferentially under most plating conditions. This behavior is typically observed in codeposition of iron-group metals (namely, Fe, Co, and Ni) or in codeposition of an iron-group metal with Zn or Cd, with either inhibition or acceleration of the rate of deposition of one of the alloying elements by the other. The term anomalous codeposition was first introduced by Abner Brenner in 1963. A well-known example of anomalous

alloy deposition is the plating of Permalloy™, a Ni–Fe alloy most commonly formed with a chemical composition of 80 at.% Ni and 20 at.% Fe, which has high magnetic permeability and high electrical resistance. The standard potentials of Ni and Fe are −0.250 and −0.440 V versus SHE, respectively. Hence, it could be expected that equal concentrations of the two metals in solution would yield a higher concentration of Ni in the deposited alloy, and it would be easy to reach the desired atomic ratio of 80 : 20. In practice, quite the opposite is observed. It seems that the presence of Fe^{2+} ions in solution reduces the rate of deposition of Ni, so that a large excess of Ni^{2+} is needed to obtain the high Ni:Fe ratio required for this alloy. This type of anomalous behavior is frequently encountered in alloy plating, in different variations. In some cases, one can observe synergism of both metals – adding one metal increases the partial current density for the depositing of the other and vice versa. Sometimes, mutual inhibition is observed, but sometimes it is a sort of one-way relationship, as in the case of Permalloy – adding Fe^{2+} to a solution containing Ni^{2+} slows down the rate of Ni deposition, but adding Ni^{2+} to a solution of Fe^{2+} increases the rate of Fe deposition.

Last, but not least, metal deposition presents an interesting challenge to the understanding of the mechanism of charge transfer across the metal/solution interphase. Alloy deposition also presents special challenges in understanding the details of processes taking place at and across the interphase. There are only a limited number of cases (for example, Pb–Sn) for which alloy deposition is "normal," in the sense that the composition of the alloy could be predicted from the behavior of the alloying elements. In most cases (for example, Ni–Re), it is "anomalous" in the sense that the composition of the alloy cannot be predicted from the known thermodynamic and kinetic parameters of the alloying elements.

The third type of electrodeposition of alloys is *induced codeposition*, a term coined by Brenner in 1963. Certain elements, such as W, Mo, Ge, and P, cannot be deposited alone from their aqueous solutions. Nevertheless, they may readily be codeposited with iron-group metals. For example, alloys of W with the iron-group metals can readily be formed using, for example, a solution of $NiSO_4$ and Na_2WO_4, with citric acid added as a complexing agent. In this particular case, it was shown that a Ni–W alloy is deposited from an adsorbed complex containing both metals, while Ni is also deposited in parallel reactions from its complex with citrate. A very similar behavior is observed for deposition of alloys of Mo. Induced codeposition may undoubtedly be considered to be "anomalous" in the sense that the composition of the alloy cannot be predicted from the electrochemical behaviors of the individual alloying elements. Adsorbed complexes containing both metals and serving as precursors for the deposition of the alloy have been suggested as one possible mechanism of both anomalous and induced codeposition. The term induced codeposition may also be used to describe a process where a metal that can barely be deposited alone, with a low current efficiency and poor adherence of the deposit, is readily deposited in the presence of other metal ions. This is the case of Re, which can be electroplated alone at FE $\leq$ 7% and poor coating quality. By adding a suitable iron-group metal salt to the bath, coatings with a Re content as high as 93 at.% and a FE as high as 96% have been obtained by Gileadi, Eliaz, and coworkers.

18.1.13 The Mechanism of Charge Transfer in Metal Deposition

18.1.13.1 Metal Ions Crossing the Interphase Carry the Charge Across it

In Section 5.5, we discussed the mechanism of charge transfer across the metal/solution interphase. Two classes of electrode reactions were defined: (i) Outer-sphere charge-transfer processes, in which both the reactant and the product are on the solution side of the interphase and charge is carried across the interphase by the transfer of one or more electrons. (ii) Inner-sphere charge-transfer processes, in which both charge and mass cross the interphase. Metal deposition and dissolution are the most important cases of the latter. It is not obvious, at first sight, which species carries the charge in metal deposition – an electron, as in outer-sphere charge transfer, or an ion. There is no doubt, however, that these processes occur on a widely different time scale – electron transfer occurs in about 1×10^{-15} seconds (femtoseconds), while the transfer of a typical metal ion would take about 1×10^{-9} seconds (nanoseconds), namely, about 6 orders of magnitude longer. Moreover, the transfer of an electron would create a highly unstable species in solution, as in the hypothetical reaction

$$[Ag(H_2O)_n]^+_{soln} + e^-_M \rightarrow [Ag(H_2O)_n]^0_{soln} \rightarrow Ag^0_{soln} \tag{18.25}$$

This would increase the total Gibbs energy of the system, violating the law of conservation of energy. Eventually, the stabilizing energy will be provided when the metal atom in solution will reach the surface and be incorporated into the metal, but this would happen a very long time later, on the femtosecond scale of electron transfer.

It has been commonly accepted for a long time that in metal deposition and dissolution, charge is carried across the interphase by the metal ions. Unfortunately, the detailed mechanism of this process has not been discussed in most cases. In particular, it is common to treat the process of deposition of divalent metals as if it involved two discrete electron transfer steps, with an intermediate stabilized by adsorption on the surface.

Here, we shall point out only two consequences of the fact that ions carry the charge across the interphase, while the electrons are transferred from the cathode to the anode through the metal wires:

(a) In many cases (e.g. Ni^{2+}, Fe^{2+}, Cd^{2+}, Sn^{2+}, Pb^{2+}) the monovalent intermediate would be unstable, and the energy of activation that would be required to form it is expected to be very high.

(b) The instability of the monovalent intermediate has not escaped the attention of people who proposed a mechanism of consecutive transfer of two electrons. The way to justify it was to assume that the monovalent intermediate would be stabilized by adsorption on the surface. However, the first electron transfer for deposition of, say, Fe^{2+}, would have to create a highly unstable intermediate in solution, following the reaction

$$[Fe(H_2O)_n]^{2+}_{soln} + e^-_M \rightarrow [Fe(H_2O)_n]^+_{soln} \tag{18.26}$$

leaving the highly unstable monovalent ion in solution, just as in the case presented by Eq. (18.1).

(c) The Butler–Volmer formalism for determining the value of the Tafel slope for different mechanisms, assuming that the symmetry factor, β, is (approximately) 0.5, was developed for electron transfer, and there is no a priori reason to assume that it would apply, or even that it would be a good approximation, for the case of ion transfer as the mode of transferring the charge across the interphase.

18.2 Electroless Deposition of Metals

18.2.1 Some Fundamental Aspects of Electroless Plating of Metals and Alloys

Electroless plating is a process in which a catalytic substrate is immersed into an aqueous solution containing a soluble salt of the metal or metals being deposited and a suitable reducing agent. It is based on chemical reduction reactions and does not require an external potential. The electrons required for the metal reduction are supplied by the simultaneous oxidation of a reducing agent. The process is self-initiating and autocatalytic.

Electroless deposition has been known for a long time. One of its early uses was the deposition of a mirror-like layer of Ag on the internal surfaces of Dewar flasks for improved thermal isolation, and as the back coating of mirrors. Already in 1835, Justus von Liebig obtained metallic coating of Ag on a glass surface by means of a chemical reduction process. In 1946, Brenner and Riddell developed electroless plating of amorphous Ni films from an aqueous solution containing sodium hypophosphite reducing agent.

Over the years, electroless deposition has found a variety of industrial applications, mainly in the preparation of corrosion- and wear-resistant protective layers, as well as metallization of plastics and semiconductors. Electroless Cu, for example, is used extensively for printed wiring boards. Binary Co–P, Ni–P, and Ni–B alloys and ternary Co–Ni–P, Co–W–P, and Ni–W–P alloys have been studied as barriers and capping layers on silicon dioxide (SiO_2) for copper metallization. Electroless Co–W–P layers are already applied in commercial process. In addition, "all-wet processes" using electroless plating have been introduced for copper ULSI and solar cell metallization. Selective electroless deposition on patterned substrates has also been applied in micro/nano fabrication.

The process of electroless plating consists of the following steps:

(1) Non-catalytic substrates, such as non-catalytic metals, plastics, ceramics, and semiconductors, are activated prior to plating in order to initiate the autocatalytic deposition process. This activation is performed by generating catalytic nuclei on the surface. Removal of metal oxides and organic impurities from the surface is essential for further activation, in order to achieve good metal film adhesion.
(2) After activation, the item is immersed in the plating solution. The main components of an electroless solution are as follows: (i) A source of metal ions. (ii) A reducing agent, which supplies the electrons for deposition.

Several reducing agents are commonly used, including hypophosphite, amine-boranes, sodium borohydride, formaldehyde, and hydrazine. (iii) Complexing agents. These are organic acids or their salts, added to form metal complexes in solution. The addition of a complexing agent prevents the spontaneous reaction in the solution, and thus controls the reaction so as to take place only on the catalytic substrate. The rate of metal deposition is proportional to the rate at which the metal complex dissociates to form free metal ions. Thus, the deposition rate is slowed down and the quality of the deposited films is improved. The reducing agent enables the deposition of neutral metal atoms on the catalytic substrate, forming a continuous film. It is important for deposition to continue when the substrate is completely covered; thus, the coated metal must maintain a catalyzed surface.

18.2.2 The Activation Process

Catalytic nuclei of metal catalyst M on a non-catalytic surface can be formed by an oxidation–reduction reaction:

$$M^{z+} + \text{Red} \rightarrow M + \text{Ox} \tag{18.27}$$

The classical catalyst for electroless deposition is Pd. The nucleation agent, M^{Z+}, in this case is Pd^{2+} (e.g. from $PdCl_2$), and the preferred reducing agent, Red, is Sn^{2+} (e.g. from $SnCl_2$). The overall reaction is thus

$$Pd^{2+} + Sn^{2+} \rightarrow Pd^0 + Sn^{4+} \tag{18.28}$$

Sn^{2+} can reduce the Pd^{2+} ions since the standard oxidation–reduction potential of Sn^{2+}/Sn^{4+} is +0.150 V and that of Pd^{2+}/Pd^0 is +0.987 V versus SHE. The flow of electrons is from the more electronegative couple (Sn^{2+}/Sn^{4+}) toward the less electronegative one (Pd^{2+}/Pd^0).

Other activation procedures have been also proposed: displacement deposition, reduction of adsorbed ions on non-catalytic substrate by reducing agent, photochemical activation that is based on photoelectron chemical kinetics, and formation of nanoparticles with high catalytic activity. Evidently, the surface coverage by the catalyst and the surface catalytic activity have a major influence on the uniformity, roughness, and adhesion to the substrate of the electroless deposited film.

18.2.3 The Reducing Agent

The composition of the electroless plating bath should be chosen such that the rate of metal deposition will be well below the mass-transport-limited rate. This can be achieved by proper selection of the type and concentration of the reducing agent.

A limited number of reducing agents are used in electroless deposition. The reduction potential (also known as redox potential) of the reducing agent must be less noble (i.e. more negative) than the reversible potential of the deposited metal. In order to reduce base metals, the redox potential of the reducing agent must be in the range of ca. −0.5 to −1.6 V versus SHE. Such reducing agents include

sodium hypophosphite ($NaH_2PO_2 \cdot H_2O$), formaldehyde (HCHO), sodium borohydride ($NaBH_4$), dimethylamine borane ($(CH_3)_2NHBH_3$), and hydrazine (H_2NNH_2).

Hypophosphite baths are the most common ones for electroless deposition of Ni- and Co-based alloys. In comparison to hypophosphites, boron-containing reducing agents allow deposition at relatively low temperatures (in the range of 40–50 °C). Sodium borohydride can be used only in strong alkaline baths. Amineboranes can be used in both acidic and alkaline baths. Some possible electron-generating chemical reactions for hypophosphites, borohydrides, and amine boranes are given in Eqs. (18.29)–(18.31), respectively:

$$H_2PO_2^- + H_2O \rightarrow H_2PO_3^- + 2H^+ + 2e^- \tag{18.29}$$

$$BH_4^- + 8OH^- \rightarrow B(OH)_4^- + 4H_2O + 8e^- \tag{18.30}$$

$$R_2NHBH_3 + 3HOH + OH^- \rightarrow R_2H_2N^+ + B(OH)_4^- + 5H^+ + 6e^- \tag{18.31}$$

Since these reactions occur in parallel to metal deposition, P and B are often incorporated in the deposited layer in an amount dependent on deposition conditions such as pH, temperature, and bath composition. Evidently, the choice of the active reducing agent, its concentration, etc. affect the incorporation of P and B in the deposited alloy and its properties. For example, heat-treated Ni–B alloy films have higher hardness and better wear resistance compared to pure Ni. Ni–P coatings have better corrosion resistance than pure Ni due to the formation of an amorphous phase. When a pure metal is required, hydrazine should best be used as the reducing agent. The oxidation reaction in this case is

$$N_2H_4 + 4OH^- \rightarrow N_2 + 4H_2O + 4e^- \tag{18.32}$$

18.2.4 The Complexing Agent

Addition of a complexing agent to the electroless plating bath is essential in order to control the reaction so that it occurs only on the catalytic surface. Complexing agents are organic acids or their salts. The main goal of the complexing agents is to control the number of free electrons available for reaction by forming metal-complex compounds and, thus, lowering the rate of the reduction reactions. Moreover, the complexing agent may have a buffering action that prevents fast variations in pH. Complexing agents also affect the quality of the deposit, in particular the evolution of internal stresses and porosity.

18.2.5 The Mechanism of Electroless Deposition

In electroless deposition, the reducing agent in solution, Red_{soln}, is the electron source; metal reduction is possible due to the simultaneous oxidation of the reducing agent. The two reactions occur on the same electrode (catalytic substrate), at the electrode/electrolyte interphase. The overall reaction can be written as

$$M_{soln}^{Z+} + Red_{soln} \xrightarrow{\text{catalytic surface}} M_{lattice} + Ox_{soln} \tag{18.33}$$

where Ox is the oxidation product of the reducing agent. It should be noted that the reaction represented by Eq. (18.33) must be conducted in such a way that a homogeneous reaction between M^{Z+} and Red in the bulk of the solution is suppressed.

The electrochemical model for electroless deposition was suggested by Paunovic and Saito on the basis of the Mixed-Potential Theory for corrosion processes (see Section 17.4.4). The oxidation process is possible if its reversible potential, $E_{rev,Red}$, is more negative than that of metal deposition, $E_{rev,M}$. The reduction and oxidation reactions attempt to reach an equilibrium potential, E_{eq}. The result of this process is the establishment of a steady state in which the mixed potential, E_{mp}, is obtained. The mixed potential corresponds to two different overpotentials:

$$\eta_M = E_{mp} - E_{rev,M} \tag{18.34}$$

$$\eta_{Red} = E_{mp} - E_{eq,Red} \tag{18.35}$$

At the steady-state mixed potential, the rate of reduction, given by the cathodic current density j_c, equals the rate of oxidation of the reducing agent, given by the anodic current density, j_a. Thus, the deposition rate is equal to

$$j_{dep} = j_c = j_a \tag{18.36}$$

The mixed potential and the deposition rate can be obtained from the current–potential curves of the partial reactions. The Wagner–Traud diagram shown in Figure 18.6 represents the general electroless metal deposition process in terms of current–potential curves. The bottom curve represents the reduction of M^{z+} ions in the absence of a reducing agent (if the activity of M^{z+} equals 1, $E_{rev,M} = E^0_M$). Similarly, the upper curve represents the oxidation of the reducing agent in the absence of M^{z+} ions. The middle curve, $j_{tot} = f(E)$, is the current–potential curve for the overall reaction. The mixed potential, E_{mp}, is the potential at which the total current, I_{tot}, equals zero, and $j_a = j_c$.

The current–potential curves can also be presented by Evans diagrams, such as that shown in Figure 18.7. The mixed potential, E_{mp}, and the deposition current density are at the intersection of the metal reduction line and the reducing agent oxidation line. As in the case of corrosion, the mixed potential represents steady state with two reactions involved; it is not an equilibrium state. The Tafel slopes for the anodic and cathodic reaction are not in any way related to each other, and the relationship $\alpha_a + \alpha_c = n$ does not apply (although it may happen to be numerically correct). Indeed, even the number of electrons transferred may not be the same for the two reactions. The only constraint determining the position of E_{mp} is that the total anodic current must be equal to the total cathodic current. Micro-polarization at potentials close to the mixed potential can be used to determine the current density during electroless plating, following the formalism used in Section 17.5.2.2, replacing j_{corr} by j_{dep}. This approach is better than determining the kinetic parameters of the two reactions (reduction of the metal and oxidation of the reducing agent) in separate solutions containing only one of these materials. The latter is theoretically correct, but it ignores the fact that the two components in solution may interact with each other, or influence the

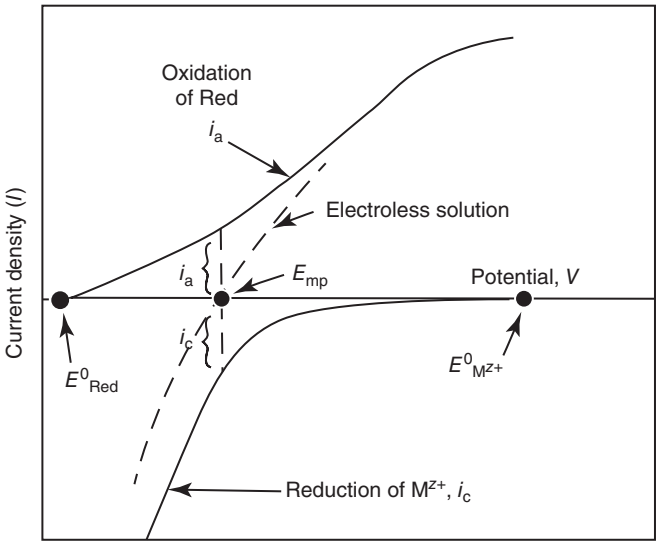


Figure 18.6 The Wagner–Traud diagram for an electroless deposition process. The partial anodic and cathodic currents are represented by i_a and i_c, respectively. The dashed line indicates the curve for the complete electroless solution.

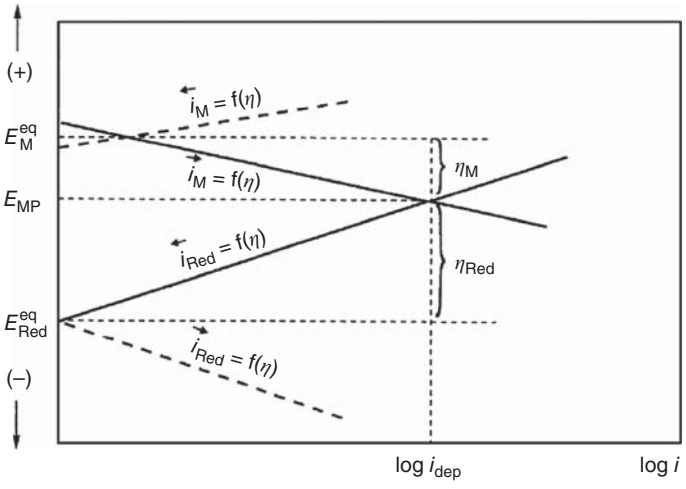


Figure 18.7 An Evans diagram for an electroless deposition process with two different simultaneous electrochemical reactions.

rate of each other in a manner similar to the processes occurring in anomalous deposition of alloys.

A superficial view would lead us to think that electroplating and electroless plating represent essentially the same process – in both cases the metal ions are being reduced. In both cases the potential is maintained by and large constant, even if the current is the externally controlled parameter, as long as the solution is not exhausted from its electroactive components. The electrons provided for the

deposition process originate in one case from an electronic device (a potentiostat or a constant-current supply) and in the other case from a reducing agent in solution, so the reduction process is the same. In view of the inherent similarity between electroless deposition and electroplating, it is not surprising that anomalous codeposition and induced codeposition are observed in both cases.

However, as it turns out, there are some important differences. In electroless plating there is only one electrode, which is the substrate being coated. In addition, electroless plating is an open-circuit process which, in a sense, is similar to corrosion; both processes occur at open circuit, so there is no current flowing in an external circuit (indeed, there is no external circuit). In corrosion, the rate of anodic oxidation of the metal must be equal to the rate of cathodic reduction of some species in solution (usually, H_3O^+, H_2O, or O_2). At steady state, a mixed potential is set up at this point. The same logic applies for electroless plating, except that the roles are reversed. Here, the rate of anodic oxidation of a suitable reducing agent must be equal to the rate of cathodic reduction of metal ions. The potential across the interface is controlled by the balance between the rate of oxidation of the reducing agent and the rate of reduction of the metal ions, either of which could be the rate-limiting factor, depending on the composition of the bath. Hence, the rate of metal deposition can be determined, in principle, by micro-polarization measurements, following the common method to determine the corrosion current.

18.2.6 Advantages and Disadvantages of Electroless Plating Compared to Electroplating

Electroless plating has two major *advantages* compared to electroplating:

1. The most important aspect of electroless deposition is that it can be performed on nonconducting surfaces, such as plastics, glasses, ceramics, and semiconductors. Although special surface treatments may be needed, a well-adherent thin metallic surface can be produced, which can be followed by electroplating of the same or a different metal. This is widely used in many engineering applications, as well as for ornamental purposes.
2. Primary current distribution, caused by the shape of the part being plated and the distance of various parts from the anode, has no effect in electroless deposition, because there is obviously no anode. Consequently, uniform metallic coating can readily be formed onto a substrate with an irregular shape and high aspect ratio features. In this sense, electroless deposition can be considered as being conducted under conditions equivalent to electrodeposition under secondary current distribution. On the other hand, partial mass-transport limitation is not eliminated; hence, ternary current distribution, which determines the micro throwing power and, in some cases, the surface morphology of the deposit, cannot be ignored. As long as the rate of deposition of the metal is low compared to the mass-transport-limited rate, the concentration at the surface is close to that in the bulk, and smooth and bright surfaces can be formed.

Electroless plating also has some *disadvantages* compared to electroplating:

1. The plating bath is inherently unstable (since it is this thermodynamic instability that allows the bath to operate). Hence, a delicate balance must be struck between the desire to have a stable bath and, yet, allow reasonably high rates of deposition.
2. The surface of the substrate being plated must be activated, to ensure that deposition will only occur where needed.
3. The rate of electroless deposition cannot be controlled as readily as the rate of electroplating, although some control can be achieved by varying the concentration of the reducing agent in the bath and by controlling the rate of mass transport and the temperature.
4. Electroless deposition is usually a slow process, and it is therefore limited to formation of relatively thin layers (typically, less than 1 μm). This could be an advantage when the purpose is to create a very thin barrier layer. Moreover, since the development of surface roughness in metal deposition has a built-in positive feedback effect, causing the roughness to increase with increasing thickness of the deposit, this is less of a problem in the case of electroless deposition. However, when a protective layer is needed, for example, to prevent corrosion or wear, this small thickness is an obvious disadvantage. This, however, is not an inherent limitation of the method, and it would seem that technologies for electroless plating of thicker layers could be developed.
5. Metal coatings formed by electroless deposition are not always pure. They might also contain phosphorus or boron, which might be derived from the reducing agent.

18.3 Electrophoretic Deposition (EPD)

EPD is a process in which charged colloidal particles dispersed or suspended in a liquid medium migrate under the influence of an electric field (*electrophoresis*) and are deposited onto an electrically conductive surface of opposite charge. The electric field is traditionally DC, although AC has also been applied in recent years. All colloidal particles that can be used to form stable suspensions and that can carry charge can be used in EPD. This includes materials such as ceramics (including carbides, oxides, nitrides, and glasses), metals, polymers, pigments, and dyes. EPD yields only a powder compact; therefore, it should be followed by a densification step, such as sintering or curing, in order to obtain a fully dense material.

EPD can be carried out both in aqueous and nonaqueous media. In certain applications, such as the deposition of ceramic materials, it is sometimes desirable to apply voltages higher than 3–4 V in order to achieve high coating thickness or to increase the rate of deposition. In aqueous solutions, such voltages would cause electrolysis of water and oxygen evolution. To avoid it, organic solvents are used instead of water as the liquid medium. These are generally polar solvents such as alcohols (e.g. ethanol) and ketones (e.g. acetone and methyl ethyl ketone).

There are two types of EPD, depending on which electrode the deposition occurs on. When the particles are positively charged, the deposition occurs on

the cathode, and the process is called *cathodic EPD*. In contrast, the deposition of negatively charged particles on the positive electrode (anode) is called *anodic EPD*.

EPD processes have several *advantages* that make them attractive for industrial use:

1. Coatings with uniform thickness can be obtained. A wide range of thicknesses, from less than 1 μm to more than 100 μm, is achievable.
2. Relatively high purity of the coated material.
3. Easy control of the coating composition, thickness, and morphology.
4. Applicability to wide range of materials.
5. Highly efficient utilization of the coating materials results in lower costs relative to other processes.
6. Complex objects can easily be coated, both inside cavities and on the outside surfaces.
7. Relatively high deposition rate.
8. The process is normally automated, needs simple apparatus, and requires less human labor than other coating processes.
9. Compared to other advanced shaping techniques, the EPD process is very versatile since it can be modified easily for a specific application. For example, deposition can be made on flat, cylindrical, or any other shaped substrate with only minor change in electrode design and positioning.
10. The aqueous process, which is commonly used, has less risk of fire relative to the solvent-borne coatings that they have replaced.
11. Modern electrophoretic paint products are significantly more environmentally friendly than many other painting technologies. There is no requirement for binder burnout as the green coating contains few or no organics.

But, EPD also suffers from some *drawbacks*, including low density, cracking of the coating due to post-deposition sintering, low adhesion strength, and need for a conductive substrate.

EPD can be used for producing coatings, for shaping monolithic, laminated, and graded free-standing objects, and for infiltration of porous materials and woven fiber preforms in the production of composites. EPD has been used industrially to coat automobile bodies and parts, heavy equipment, appliances, metal furniture, beverage containers, fasteners, phosphor coatings in the manufacture of screens for cathode ray tubes (CRTs) for advanced display applications, etc. EPD is used to fabricate porous ZrO_2 anodes for solid oxide fuel cells, functional films for advanced microelectronic devices, novel composites or bioactive coatings for medical implants, nanosize zeolite membranes, glass and ceramic matrix composites by infiltration of ceramic particles onto fiber fabrics, functionally graded ceramics, superconductors, piezoelectric materials, etc. Since the 1920s, electrophoretic painting has been used for the deposition of rubber latex.

A distinction should be made between electroplating and EPD: (i) In electroplating the moving species are ions, whereas in EPD they are solid particles suspended in a solvent. (ii) Electroplating involves charge transfer in the form of ion reduction, whereas EPD involves no charge transfer. (iii) In electroplating the liquid medium should better have high conductance, whereas in EPD low

conductance is required. (iv) In electroplating the liquid most commonly used is water, whereas in EPD it is organic. (v) The typical deposition rate in EPD is approximately 4 orders of magnitude larger than in electroplating.

It is essential that the particles in suspension carry charge in order for them to move in response to the electric field. Four mechanisms have been proposed for the development of charge on the particles: (i) Selective adsorption of ions onto the solid particle in the liquid. (ii) Dissociation of ions from the solid phase into the liquid. (iii) Adsorption or orientation of dipolar molecules at the particle surface. (iv) Electron transfer between the solid and liquid phases due to differences in work function.

The *zeta potential*, ζ, of particles is a key factor in the EPD process. This term was only mentioned in Section 8.1.4 and in Figure 8.5c therein, when describing the Stern model for the structure of the electrical double layer. A more comprehensive description is given here. A charged particle in a suspension is surrounded by ions with an opposite charge in a concentration higher than the bulk concentration of these ions; this is the so-called "double layer." When an electric field is applied, these ions and the particle should move in opposite directions. However, the ions are also attracted by the particle and, consequently, a fraction of the ions surrounding the particle will not move in the opposite direction but move along with the particle. Hence, the speed of a particle is not determined by the surface charge but by the net charge enclosed in the liquid sphere, which moves along with the particle. The electric potential will vary, from Ψ_0 at the surface of the particle to zero far from the particle. Between these two end points, the potential at the surface of hydrodynamic shear, or *slipping plane*, which separates between the liquid that moves with the particle and the liquid that does not move with it, is termed the zeta potential, and is the main parameter determining the electrokinetic behavior of the particle. The zeta potential is a key indicator of the stability of colloidal dispersions. Its magnitude indicates the degree of electrostatic repulsion between adjacent, similarly charged particles in a dispersion. For molecules and particles that are small enough, a high zeta potential (more positive than +30 mV or more negative than −30 mV) will confer stability, i.e. the solution or dispersion will resist aggregation. When the potential is small, attractive forces may exceed this repulsion, and the dispersion might break and flocculate. A zeta potential versus pH curve will be positive at low pH, and lower or negative at high pH. There may be a point where the plot passes through zero zeta potential. This point is called the *isoelectric point*, which is very important from a practical consideration – it is normally the point where the colloidal system is least stable.

It is imperative to achieve a high and uniform surface charge of the suspended particles. It plays a role in (i) stabilization of the suspension by determining the intensity of repulsive interaction between particles, (ii) determining the direction and migration velocity of particle during EPD, and (iii) determining the green density of the deposit. During formation of the deposit, the particles become closer to each other, and with increasing attraction force. If the charge on the particles is low, they would coagulate even for relatively large interparticle distances, leading to porous, sponge-like deposits. On the other hand, if the particles have a high surface charge during deposition, they will repulse each other, occupying positions that will lead to high particle packing density. The zeta potential can be

controlled by adding to the suspension a variety of charging agents, such as acids, bases, and specifically adsorbed ions or polyelectrolytes.

Measurement of the zeta potential is important to many industries, such as water treatment, mineral processing, pharmaceuticals, and additives for electronics. Zeta potential is not a directly measurable parameter; it should be calculated from other experimentally measurable properties of the system, using appropriate theory. ISO standards 13099 Parts 1, 2, and 3 describe electroacoustic and electrokinetic phenomena, optical methods for zeta potential determination, and acoustic methods for zeta potential determination, respectively. All zeta potential analyzers have means for moving liquid relative to the solid surface for disturbing the double layer (e.g. by external electric field or mechanical pressure field), as well as means for monitoring the generated signals (by optical, electrical, or mechanical means). Most of them are based on application of electrical fields on the particle dispersion, causing electrophoretic migration, and calculating the amount of Doppler shift that results from scattering light during migration.

Further Reading

1 Gu, Y., Liu, J., Qu, S. et al. (2017). Electrodeposition of alloys and compounds from high-temperature molten salts. *J. Alloys Compd.* 690: 228–238.

2 Duhin, A., Inberg, A., Eliaz, N., and Gileadi, E. (2015). Electroless plating of rhenium-based alloys with nickel, cobalt and iron. *Electrochim. Acta* 174: 660–666.

3 Berkh, O., Eliaz, N., and Gileadi, E. (2014). The initial stages of electrodeposition of Re–Ni alloys. *J. Electrochem. Soc.* 161: D219–D226.

4 Dickerson, J.H. and Boccaccini, A.R. (ed.) (2012). *Electrophoretic Deposition of Nanomaterials*. NY, USA: Springer-Verlag.

5 Schlesinger, M. and Paunovic, M. (ed.) (2010). *Modern Electroplating*, 5th edn. NJ, USA: Wiley.

6 Naor, A., Eliaz, N., and Gileadi, E. (2010). Electrodeposition of alloys of rhenium with iron-group metals from aqueous solutions. *J. Electrochem. Soc.* 157: D422–D427.

7 Eliaz, N. and Gileadi, E. (2008). Induced codeposition of alloys of tungsten, molybdenum and rhenium with transition metals. In: *Modern Aspects of Electrochemistry*, vol. 42 (ed. C.G. Vayenas, R.E. White and M.E. Gamboa-Aldeco), 191–301. New York, USA: Springer.

8 Kossoy, E., Khoptiar, Y., Cytermann, C. et al. (2008). The use of SIMS in quality control and failure analysis of electrodeposited items inspected for hydrogen effects. *Corros. Sci.* 50: 1481–1491.

9 Besra, L. and Liu, M. (2007). A review on fundamentals and applications of electrophoretic deposition (EPD). *Prog. Mater. Sci.* 52: 1–61.

10 Paunovic, M. and Schlesinger, M. (2006). *Fundamentals of Electrochemical Deposition*, 2nd edn. NJ, USA: Wiley.

11 Eliaz, N., Sridhar, T.M., and Gileadi, E. (2005). Synthesis and characterization of nickel–tungsten alloys by electrodeposition. *Electrochim. Acta* 50: 2893–2904.

12 Watanabe, T. (2004). *Nano-Plating: Microstructure Control Theory of Plated Film and Data Base of Plated Film Microstructure*. Oxford, UK: Elsevier.

13 Kanani, N. (2004). *Electroplating: Basic Principles, Processes and Practice*. Oxford, UK: Elsevier.

14 Budevski, E., Staikov, G., and Lorenz, W.J. (1996). *Electrochemical Phase Formation and Growth: An Introduction to the Initial Stages of Metal Deposition*. Weinheim: Wiley-VCH.

15 Dini, J.W. (1993). *Electrodeposition: The Materials Science of Coatings and Substrates*. NJ, USA: Noyes Publications.

16 Mallory, G.O. and Hajdu, J.B. (ed.) (1990). *Electroless Plating: Fundamentals and Applications*, American Electroplaters and Surface Finishers Society. NY, USA: Noyes Publications.

17 Matiašovsky, K., Fellner, P., Chrenkov-Paučrov, M. et al. (1987). Electrolytic metal coating in molten salts. *Chem. Pap.* 41: 527–565.

18 Bindra, P. and Roldan, J. (1987). Mechanisms of electroless metal plating. III. Mixed potential theory and the interdependence of partial reactions. *J. Appl. Electrochem.* 17: 1254–1266.

19 Puippe, J.C. and Leaman, F. (ed.) (1986). *Theory and Practice of Pulse-Plating*. Florida, USA: AESF.

20 Safranek, W.H. (1986). *The Properties of Electrodeposited Metals and Alloys*, 2nd edn. Florida, USA: AESF.

21 Ibl, N. (1980). Some theoretical aspects of pulse electrolysis. *Surf. Technol.* 10: 81–104.

22 Brenner, A. (1963). *Electrodeposition of Alloys*, vol. I & II. NY, USA: Academic Press.

19

Electrochemical Nanotechnology

19.1 Introduction

The emergence of nanoscience and nanotechnology has led to new developments in electrochemistry and its applications. Examples include surface biofunctionalization or coating of implants, nano-sized electrodes and capillaries and their use in biosensors or electrochemical printing (EcP), electrochemistry of freely diffusing single molecules, electrochemical scanning probe techniques and their use in the investigation of immobilized biomolecules, biomimetic membranes, etc. These have been reviewed in volumes 52 and 53 of *Modern Aspects of Electrochemistry* edited by one of us (N.E.). Here, we shall limit our discussion just to some fundamental aspects of the physics of nanoparticles (NPs) that are relevant to the preparation of electrode materials, e.g. in electrochemical energy-conversion devices, and to nano-sized electrodes and capillaries and their use in EcP. The many ingenious methods of preparation and applications of NPs in electrochemistry are widely discussed in the literature and will not be treated here.

19.2 Nanoparticles and Catalysis

Aggregates of atoms may be classified, based on the number of atoms forming them, as (i) small clusters, the structure and properties of which do not vary in a monotonic manner; (ii) large clusters; and (iii) NPs for which as their size is increased the properties slowly approach the bulk value. Clusters are aggregates of atoms, molecules, or ions, usually with less than a few thousand atoms. At this level, there are not enough atoms to classify the collection as a crystal. When the clusters become sufficiently large, they may adopt a different crystal structure as compared to the stable bulk crystal structure.

NPs constitute a major class of nanomaterials. NPs are zero-dimensional, i.e. they have nanometric sizes in all three dimensions. The diameters of NPs can vary anywhere between one and a few hundred nanometers. Small NPs with diameters of a few nanometers are comparable to molecules. Accordingly, the electronic and atomic structures of such small NPs have unusual features, markedly different from those of the bulk materials. Large NPs (>20–50 nm), on the other hand,

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

may have properties similar to those of the bulk. NPs can be amorphous or crystalline. NPs of metals, chalcogenides, nitrides, and oxides are often single crystals. Crystalline NPs are referred to as nanocrystals.

Already in 1857 Michael Faraday noted that colloidal metal sols (i.e. dispersions of solids in liquids) were thermodynamically unstable, and that the individual particles had to be stabilized kinetically against aggregation. In 1905, Albert Einstein related the Brownian motion (i.e. the random motion of particles suspended in a fluid, whether a liquid or a gas, resulting from their collision with the fast-moving molecules in the fluid) executed by NPs to their diffusion coefficient. Following Richard Feynman's two seminal lectures (1959 and 1984) that introduced the world to nanotechnology, there has been a great upsurge in the use of colloid chemical methods to generate NPs of different materials.

19.2.1 Surfaces and Interfaces

An *interface* is a small number of molecular layers that separate between two adjacent bulk condensed phases. The *surface* of a solid is a particularly simple type of interface, at which the solid is in contact with its surrounding environment. The surface of an object determines its chemical reactivity, wetting behavior, optical appearance, frictional behavior, etc. The interface between two phases has unique properties that are different from the bulk properties of either phase in contact. This gives rise, among other things, to the double-layer capacitance, C_{dl} (see Chapter 8).

While atoms in the bulk are symmetrically coordinated, on the surface they lack that symmetry and, to a first approximation, have only half as many atoms within range of significant interaction. In addition, a typical material will have a volume density of $\sim 10^{23}$ atoms cm^{-3} and a surface density of only $\sim 10^{15}$ atoms cm^{-2}.

Molecules in the bulk of a material are at a lower energy state than at the surface due to nearest-neighbor interactions. Hence, increasing the surface area requires added energy, given by

$$\mathrm{d}G = \gamma\,\mathrm{d}A \tag{19.1}$$

where $\mathrm{d}A$ represents the incremental increase in surface area that leads to a corresponding increase of $\mathrm{d}G$ in the Gibbs free energy. γ is the *surface energy*, i.e. the reversible work required to create a unit area of surface at constant volume, temperature, and chemical potential of the species. It is defined in units of energy per unit surface area ($\mathrm{J\,m^{-2}}$ or $\mathrm{N\,m^{-1}}$). For liquids, γ is numerically equal to the surface tension, σ. In most cases, this is not true for solids. The surface tension also acts to decrease the free energy of the system. It is defined in units of force per unit length ($\mathrm{N\,m^{-1}}$), i.e. the same units as those of γ. Although dimensionally γ and σ are identical, physically they are different types of quantities; while γ is a scalar, σ is a second-order tensor. Below a critical temperature, a liquid will minimize its free energy by seeking a shape that exposes the least possible number of atoms or molecules at its surface. Consequently, water forms roughly spherical droplets, and meniscus effects are observed in capillaries.

Dealing with interfaces and surfaces, it should be borne in mind that they readily contaminate, have better ability to form bonds with a reactant

or intermediate in solution, and have a larger Gibbs energy of adsorption. Moreover, surface diffusion is faster than diffusion in the bulk material. In addition, traditional techniques used to characterize the bulk structure of materials are often not suitable for surface analysis.

While in large objects with small surface area to volume ratio (A_S/V) the physical and chemical properties are primarily defined by the bulk, in small objects with a large A_S/V ratio the properties are strongly influenced by the surface. The surface is particularly important in the case of NPs. It is possible to estimate the fraction of atoms on the surface of the particle, P_S, using the simple relation

$$P_S = 4N^{-1/3} \tag{19.2}$$

where N is the total number of atoms in the particle. The fraction of surface atoms becomes less than 1% only when the total number of atoms is of the order of 10^7, which for a typical metal would correspond to a particle diameter of 150 nm.

19.2.2 The Vapor Pressure of Small Droplets and the Melting Point of Solid NPs

The increase of the Gibbs energy of a small droplet of liquid with decreasing radius results in an increase in vapor pressure, expressed by the Kelvin equation

$$2.3RT \log\left(\frac{p_r}{p_0}\right) = \frac{2\gamma}{r}\frac{M}{\rho} \tag{19.3}$$

where M is the atomic mass and ρ is the density. The ratio p_r/p_0 is the vapor pressure of a drop of radius r divided by that of a flat surface at the same temperature. A plot of p_r/p_0 versus the radius is shown in Figure 19.1a for several liquids. Considering that the boiling point is defined as the temperature at which the vapor pressure reaches one atmosphere, it follows that a small droplet of liquid will boil at a lower temperature than bulk liquid. In Figure 19.1, the curves for water and ethanol are almost the same, because the ratio $\gamma M/\rho$ happens to be the same, although the three terms in this ratio are quite different.

When a solid is involved, the Kelvin equation cannot be used because the surface tension cannot be readily measured and the NPs of metals are not necessarily spherical. Moreover, mechanical stress that may exist in very small particles could influence the Gibbs energy of adsorption. Nevertheless, the physical rules are the same, and the excess surface Gibbs energy has been found to lower the melting point of the metal.

We can calculate the percentage of atoms on the surface as a function of the size of the particle. To do this properly, one should know the shape of the particle, which could be different for different metals. For simplicity, we shall assume a spherical particle. The volume of a spherical shell is given by $4\pi r^2 \Delta r$, where Δr is the thickness of the shell. Hence, the percentage of atoms on the surface of a sphere is given by

$$\frac{4\pi r^2 \Delta r}{(4\pi/3)r^3} \times 100 = \frac{3\Delta r}{r} \times 100 \tag{19.4}$$

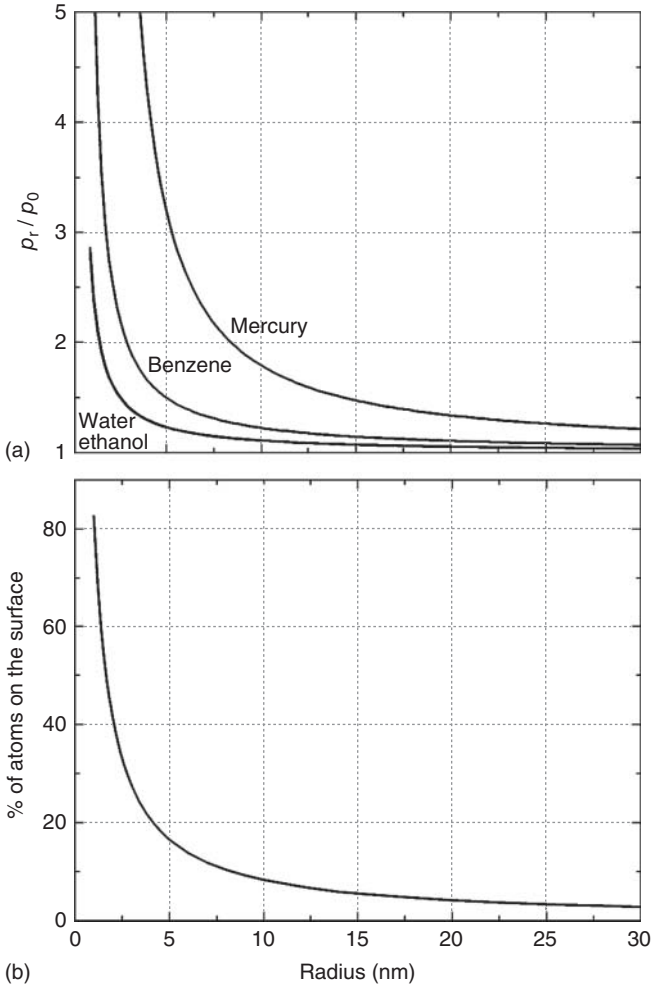


Figure 19.1 (a) The increase in vapor pressure of small drops of different liquids with decreasing radius of the drop at 25 °C. (b) The percentage of atoms on the surface of a Pt nanoparticle, as a function of its radius.

For a spherical NP of Pt, the value of Δr was estimated on the basis of the number of atoms per cm^2 on the surface of Pt, known from the measured value of the maximum surface concentration, Γ_{max}, for hydrogen adsorbed on Pt, which is 1.3×10^{15} atoms cm^{-2}, yielding a thickness of $\Delta r = 0.277$ nm. Note that this is the thickness of the spherical shell corresponding to the diameter of a Pt atom. The result is shown in Figure 19.1b. The curves shown in Figure 19.1 are quite similar, as may be expected, since both are related to the special activity of the surface of an NP.

The effect of size becomes significant for $r \leq 10$ nm, namely, where NPs are concerned. In contrast, for a typical colloidal particle ($r \approx 1$ μm), the extrapolated value of the curves in Figure 19.1 both yield an insignificant deviation of about

0.1% from bulk properties. It is interesting to note that, small as it may seem, the radius of 10 nm is rather large on the atomic scale. For example, a water droplet of this radius contains about 1.4×10^5 molecules, and a similar sized Pt sphere would contain about 2.8×10^5 atoms.

The effect of particle size on the melting points of Al and Pb is shown in Figure 19.2, respectively. The bulk melting point of Al (660.3 °C) is reduced to about 565 °C for a radius of 6.7 nm. It is noted also that the effect of size on the melting point becomes important only when the radius is somewhat below 15 nm, which is consistent with the calculations given in Figure 19.1. A similar (but by no means identical) behavior was observed for Pb, as shown in Figure 19.2b. In this case, the melting point does not deviate from the bulk value until $r < 7.5$ nm, but then it is lowered sharply from 327.5 °C to about 133 °C as the radius is reduced to 1.4 nm.

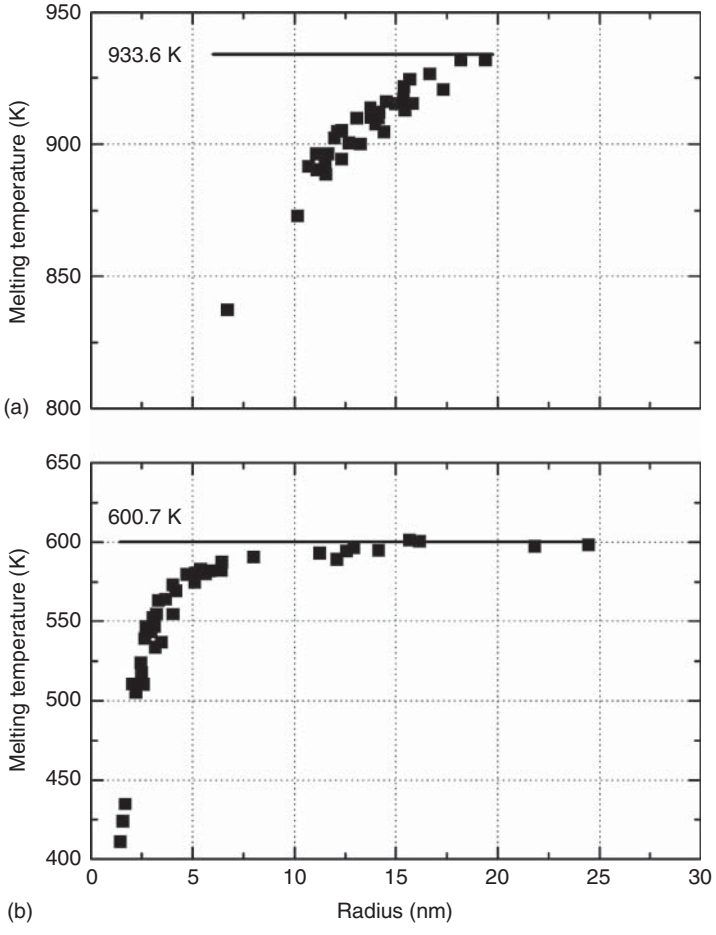


Figure 19.2 The melting point of (a) Al and (b) Pb, as a function of the radius of the NP. Source: Reprinted with permission from D. Vollath, *Nanomaterials – An Introduction to Synthesis, Properties and Applications*, Wiley-VCH, Weinheim, Germany, 2008.

19.2.3 The Thermodynamic Stability and Thermal Mobility of NPs

An isolated NP in the bulk of a solution or on the surface of a suitable catalyst can be stable, in spite of its high surface energy, as reflected by the high vapor pressure (Figure 19.1a) or lowering of the melting point (Figure 19.2). However, when two such particles collide and are fused together, their excess surface free energy per mole decreases. Assuming, for simplicity, that the particles have an equal radius, the radius of the new particle formed (assumed to be spherical) will increase by a factor of $\sqrt[3]{2} = 1.26$. The position of the new particle on the lines in Figure 19.1 will move to the right, lowering the vapor pressure or the melting point (and, hence, the excess surface Gibbs energy). Put differently, it is correct to say that a NP is thermodynamically unstable with respect to a process of merging it with one or more NPs. This simple conclusion is very important in the area of catalysis or electrocatalysis, because it could lead to sintering of the catalytic NPs into lumps of the same material, losing part of their active surface area as well as any inherent catalytic activity they may have had as NPs.

Being "thermodynamically unstable" is not an insurmountable impediment in technology. Indeed, almost everything we use has this property: the wooden furniture in our houses is flammable, being unstable in respect to reaction with oxygen in the air. Nevertheless, furniture can survive for centuries without showing any sign of burning slowly. The cutlery we use is often made of stainless steel – an alloy of iron, chromium, and nickel, all of which can combine with oxygen to form various oxides, lowering the Gibbs energy. Still, cutlery can be inherited and maintained for generations, showing little corrosion. This *apparent stability* is imparted by slow kinetics. Wood may dry out, but it cannot ignite and burn spontaneously. Cars are manufactured with protective coatings to minimize corrosion, and metals such as aluminum, titanium, and niobium form their own protective oxide films in air spontaneously. All of these lower the rate of corrosion, the reaction leading down the Gibbs energy slope toward equilibrium. The stability may be *apparent* from the thermodynamic point of view, but it can be very real from the engineering point of view.

19.2.4 Catalysts

A catalyst may be defined as (i) a substance that enables a chemical reaction to proceed at a usually faster rate or under different conditions (e.g. at a lower temperature) than otherwise possible; or (ii) an agent that provokes or speeds up significant change or action. There are five classes of catalysts: (i) biocatalysts, e.g. enzymes. These catalysts operate at ambient temperature and pressure in an aqueous liquid. Water oxidation in photosynthesis is an example of a reaction affected by biocatalysts. (ii) Homogeneous catalysts, e.g. transition-metal complexes. These catalysts are required to operate at low-to-intermediate temperatures, and a wide range of pressures, in a liquid (either aqueous or non-aqueous). Olefin metathesis is an example of a reaction that uses such catalysts. (iii) Heterogeneous catalysts, e.g. Rh NPs. These catalysts are required to operate at a wide range of temperatures and pressures, either in gas or liquid phase. NH_3 synthesis is an example of a reaction that uses such catalysts. (iv) Electrocatalysts,

e.g. Pt NPs. These catalysts are required to operate at a wide range of temperatures and pressures, either in gas or liquid phase. They are used, for example, for water electrolysis in fuel cells. Electrocatalysts come in different forms: photocatalysts, solution-phase electrocatalysts, and surface electrocatalysts. (v) ultrahigh vacuum (UHV) catalysts, e.g. Ru. They are used in fundamental studies in surface science, for example, of adsorption–desorption reactions. They are operated at very low pressure (vacuum) within a wide range of temperatures.

Catalysts play an important role in reactions such as oxidation of hydrogen, hydrogenation of hydrocarbons, oxidation of CO, propylene epoxidation, etc. The three primary figures of merit for catalysts are activity, stability, and selectivity. The increasing proportion of surface atoms with decreasing particle size compared with bulk metals gives metal NPs unique properties. For example, NPs have better catalytic properties due to (i) the increased surface area per unit volume, and (ii) the increased surface activity due to higher degree of unsaturated bonds. Of course, the chemistry and structure of the material is also very important. In practice, there are only 12 metals (belonging to groups VIII and 1B) that are elemental catalysts, including Fe, Co, Ni, Cu, Pt, Rh, Pd, and Ag. Gold atoms, on the other hand, have a high ionization potential of 9.22 eV, thus giving it a poor affinity for molecular hydrogen and oxygen. Interestingly, Au displays a drastic change in behavior on reduction of size and can perform the role of a catalyst when it is in the nano form.

19.2.5 The Effect of Particle Size on Catalytic Activity

Considering that the surface atoms have a higher Gibbs energy than atoms in the bulk, the common wisdom is that the Gibbs energy of adsorption would be higher on the surface of NPs than on large particles. While this makes sense, it is not self-evident that a higher energy of adsorption leads to higher catalytic activity. In Section 7.5 and Figure 7.3 we discussed the so-called *volcano plot*, where the best catalyst is one that gives rise to an intermediate value of the bonding energy for adsorption. On one hand, the energy of adsorption must be high enough to attach the reactant to the surface; on the other hand, it should be low enough to allow it to react there and release the product to the solution. Different volcano plots have been presented for different electrochemical reactions (Z.W. Seh et al., 2017). Considering the volcano plot, it all depends on whether the specific rate constant, or the exchange current density that is proportional to it, is on the ascending or the descending branch of the plot of reaction rate versus the bond energy shown in Figure 7.3. If, as a result of the small size of the particle, the Gibbs energy of adsorption becomes too high, the intermediates formed in the reaction sequence could be too strongly adsorbed on the surface, reducing its catalytic activity. Indeed, it was found in some studies of catalytic activity as a function of particle size that there is a maximum in catalytic activity for particle sizes in the range of $3 \leq r \leq 5$ nm, which is about where the fraction of the atoms on the surface starts rising sharply, and the melting points of metals start decreasing very significantly, as shown in Figures 19.1 and 19.2.

Another factor that is important in the context of the activity of NPs in electrochemical systems is that adsorption in electrochemistry is a replacement reaction

that can be described by an equation of the type

$$RH_{soln} + n(H_2O)_{ads} \rightarrow RH_{ads} + n(H_2O)_{soln} \tag{19.5}$$

where RH stands for an organic species that could be the reactant or one of the intermediates in a reaction sequence. This is referred to as *electrosorption*. It is characterized by the fact that the Gibbs energy of electrosorption is the difference between that of the energy of adsorption of RH and that of n molecules of water, as discussed in Section 12.1.

$$\Delta G_{ads} = \Delta G_{ads,RH} - n\Delta G_{ads,W} \tag{19.6}$$

This equation shows that the increase of the Gibbs energy of adsorption resulting from the decrease in size of the particle may be compensated for, or at least become much less significant, as a result of a similar increase of the Gibbs energy of adsorption of one or more water molecules, depending on the size of the species being adsorbed.

The purpose of the current section is not to claim that reduction of the size of the particles of a catalyst cannot lead to an increase of its catalytic activity, but only to point out that such an increase is not universally correct, and the effect of size on activity should be tested for each system experimentally.

19.2.6 Nanoparticles Compared to Microelectrodes

The particular behavior of a single microelectrode or an ensemble of millions of microelectrodes is discussed in Section 13.3. Since a NP is the ultimate case of an ultramicroelectrode (UME), it is appropriate to discuss some of the properties of NPs employing the equations developed for microelectrodes, in order to calculate the increased rate of diffusion toward an isolated NP and the corresponding decrease in solution resistance.

For a single NP, assumed to be spherical, the limiting current is given by

$$j_L = \frac{nFDc_b}{\delta} = \frac{nFDc_b}{r} \tag{19.7}$$

where the radius of the particle plays the role of the Nernst diffusion-layer thickness, δ, in the case of semi-infinite linear diffusion. The following numerical examples show the enhanced rate of diffusion and the associated decrease in the solution resistance for a NP of $r = 5$ nm. Taking $n = 1$, $D = 6 \times 10^{-6}$ cm^2 s^{-1}, and $c_b = 1 \times 10^{-6}$ mol cm^{-3} yields, according to Eq. (19.7), a very large limiting current density of $j_L = 1.16$ A cm^{-2} for a rather dilute solution (1.0 mM) of the reactant.

It may be added here that stirring will have no influence on the limiting current density calculated above, because the Nernst diffusion-layer thickness is $\delta \geq 5$ μm, even in vigorously stirred solutions, which is 3 orders of magnitude higher than the radius of the NP. This can be seen from the equation for the limiting current density, taking into account both stirring and the radius of the NP (see Eq. (13.36)), which is given by

$$j_L = nFDc_b\left(\frac{1}{\delta} + \frac{1}{r}\right) \tag{19.8}$$

The solution resistance (see Eq. (3.9) and Section 13.3) for the same NP is given by

$$R_S = \frac{r}{\kappa} = \frac{5 \times 10^{-7}\ \text{cm}}{0.01\ \ \text{S cm}^{-1}} = 5 \times 10^{-5}\ \Omega\text{cm}^2 \tag{19.9}$$

where a moderate specific conductivity of $\kappa = 0.01\ \text{S cm}^{-1}$ has been assumed. At a current density of $j_L = 0.1\ \text{A cm}^{-2}$, the resulting potential drop is then given by

$$jR_S = 0.1\ \text{A cm}^{-2} \times 5 \times 10^{-5}\ \Omega\text{cm}^2 = 5\ \ \mu V \tag{19.10}$$

To be exact, one should point out that the above calculation is only approximate. Thus, microelectrodes are usually made as flat discs of the desired metal, embedded in an insulator and polished to be flush with the surface. NPs, on the other hand, are usually prepared separately and attached to the surface, so their interface with the electrolyte may be in the form of a hemisphere. In any case, neither is really spherical. However, the calculation above is given to show orders of magnitude of the diffusion rate and the solution resistance, and an error by a factor of two or three does not matter.

This looks like an ideal situation for conducting electrochemical measurements at current densities far below the limiting current density, with an insignificant error caused by mass-transport limitation or by the potential drop across the solution resistance; but there is a problem. The surface area of a nanosphere of 5 nm radius is about $3 \times 10^{-12}\ \text{cm}^2$. Hence, at $0.1\ \text{A cm}^{-2}$, the total current is only 3×10^{-13} A. This is measurable, but not useful for any device. In order to build a power source (e.g. a battery or a fuel cell), one would have to pack huge numbers of NPs per unit geometrical surface area. On the other hand, when the NPs are packed close together, mass transport by diffusion is reduced to the value found for planar electrodes, since the diffusion fields of all the particles completely overlap each other. The same applies to the solution resistance, which increases to values characteristic of planar electrodes.

19.2.7 The Need for High Surface Area

It has long been realized in the field of electrochemistry that increasing the surface area increases the *effective* catalytic activity. This is a simple-minded approach to electrocatalysis, ignoring the inherent properties of the catalyst, yet it is highly effective. In the time-honored method of preparing reversible hydrogen electrodes employing platinized platinum, the real surface area of the electrode is increased by roughening it, without actually increasing the physical dimension of the electrode. In this way, the effective exchange current density can be increased by as much as 2 or 3 orders of magnitude, without changing the true catalytic activity of the surface. This approach is routinely applied successfully in the area of batteries and fuel cells. Improved performance is achieved using very high surface area materials. Methods of producing such electrodes have been developed over the years. The best substrate is based on carbon, which can produce electrodes having a specific area as high as $2.5 \times 10^7\ \text{cm}^2\ \text{g}^{-1}$. Such materials could be mixed with Pt or Pt–Rh alloys, for example, to prepare similar high-surface-area catalysts.

Although there are several methods of measuring the surface area of porous materials, there seems to be some uncertainty regarding the true area where the electrochemical process can take place. Thus, employing the Brunauer–Emmett–Teller (BET) theory, one uses the adsorption of an inert gas (usually, nitrogen), measured as a function of relative pressure, to calculate the specific surface area of materials in units of $m^2\ g^{-1}$ (ISO 9277 standard is applicable). Standard BET analysis is most often conducted at the boiling temperature of N_2 (−196 °C). The specific surface area determined through BET theory may depend on the adsorbate molecule utilized and its adsorption cross-section. It should also be borne in mind that some of the area interacting with a small gas molecule may not be accessible to the solvent or to the reactant.

Another method has been to use the formation of a layer of oxygen or hydrogen atoms, followed by electrochemical stripping. This may have some advantage over methods based on adsorption of gases, because it measures the area of the surface that is in contact with the solution. In recent years, the adsorption of CO followed by its anodic stripping has been preferred because CO is very strongly adsorbed, and the results obtained are not sensitive to the presence of impurities. Nevertheless, it could be argued that this method measures the area accessible to a very small molecule, and some of the sites may not be available for adsorption of a larger molecule, such as methanol.

Attempts have been made to determine the catalytic activity of single NPs as a function of size. This is a delicate matter, since the area of such particles is of the order of magnitude of $10^{-12}\ cm^2$, and any error in the measurement (or better, the estimation) of the surface area could lead to a major error when extrapolating to a macroscopic-sized electrode. The same applies to any background or stray current, bearing in mind that a background current of 1×10^{-13} A could lead to an error of 0.1 A cm^{-2} or more. In addition, the volume-to-surface area ratio in such measurements is extremely high, so that keeping the electrode clean enough, in order to allow a meaningful determination of its inherent catalytic activity, could be an insurmountable challenge, as discussed in Section 7.1.2.

This leaves the question of the dependence of catalytic activity on the size of the NP open to debate, at least from the point of view of the theory of electrocatalysis. On the other hand, there is no doubt that employing NPs is a valid and highly effective method for producing electrodes with high surface area, thereby increasing the catalytic activity. Whether this should be attributed to an increase in the intrinsic catalytic activity associated with the small size or just to the increase in electroactive surface area of the electrode may be of secondary importance from the practical point of view, for example, in the design of better anodes for fuel cells.

19.3 Electrochemical Printing

Many materials can be electrodeposited, including metals, alloys, conducting polymers, and semiconductors. Madden and Hunter (1996) were probably the first to introduce electrochemical three-dimensional (3D) deposition – printed nickel columns with several microns resolution. Since then, many studies have

been published in this area, each describing an improvement or a new method for electrochemical additive manufacturing (AM). Some of these methods will be briefly reviewed in Section 19.3.1.

Electrochemical 3D printing of metals has some significant advantages compared to current commercial technologies for printing metal powders: (i) it can reach submicron resolution; features with less than 100 nm radius have been demonstrated. (ii) It does not involve any significant heating; thus, no heat-affected zone-related phenomena occur, residual stresses are minimal, and no subsequent heat treatment is required. (iii) It typically does not require any support structures. (iv) The chemicals and reagents are usually cheaper. (v) Less safety issues are encountered. (vi) Less recycling issues are involved. (vii) Equipment employed is much cheaper and smaller.

19.3.1 Electrochemical Printing Processes

Here, we summarize some of the major electrochemical printing processes that have been reported in the open literature. The *electrochemical fabrication (EFAB)* technology was invented by Adam Cohen et al. as an improved LIGA[1] process, and is based on a multilayer electrodeposition and planarization of at least two metals: a structural material and a sacrificial substance. The EFAB technology is a three-step additive/subtractive process, with the ability to fabricate different structures in a layer-by-layer manner. The final metal structures were found to be fully dense and with excellent mechanical properties. The EFAB technology is capable of producing metal devices, centimeters to millimeters in size, with micron-scale features as small as 20 ± 2 μm (see Figure 19.3).

The *localized electrochemical deposition (LECD)* process utilizes a tip-directed localized electric field to induce an electrochemical reaction that can deposit metals, alloys, conducting polymers, and semiconductors onto conductive substrates based upon the principles of electroplating or electrolysis. An electric

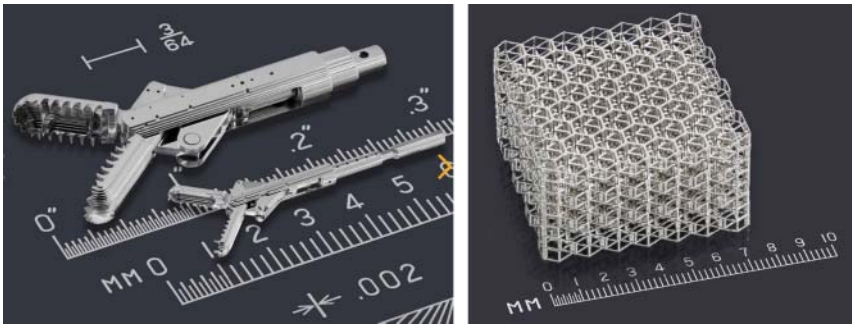


Figure 19.3 Fine objects manufactured by EFAB. Source: Reprinted with permission from http://www.microfabrica.com.

1 LIGA is a German acronym for Lithographie, Galvanoformung, Abformung (Lithography, Electroplating, and Molding) that describes a fabrication technology used to create high-aspect-ratio microstructures.

field, confined to the region beneath the microelectrode tip (e.g. Pt–Rh), will be generated when the tip of a microelectrode is brought close to the surface of the substrate at a negative electric potential relative to the tip, while both are submerged in a plating solution. Reduction of the metal ions present in the solution, and thus metal deposition, will be greatly localized within this region such that moving the microelectrode tip with respect to the forming surface can potentially create structures of any geometry. Structures are built by moving the electrode in three dimensions relative to the substrate, or vice versa. Electrode diameters between hundred micrometers and a few nanometers have been reported. Although the LECD process is considered groundbreaking, and is the first electrochemical 3D printing process, the difficulty to confine the electric field to a submicron region limits its resolution. As a result, the only features that have been successfully printed are micro-columns of polycrystalline Ni or Cu, with diameters and heights ranging from hundreds to tens of microns.

Electrochemical jet printing, also referred to as *electrochemical printing (EcP)*, is a software reconfigurable tool and process for electrodeposition of multiscale, multimaterial objects from input drawings. A potential is applied between an anode assembled in a hollow micropipette and a cathode surface. A controlled and continuous stream of ionic solution is pushed from the pipette while both the tip and the substrate are submerged in an electrolyte solution. The EcP system – comprising custom LabVIEW print driver software, hardware, electrolyte, and a microjet print nozzle – creates complex patterns by locally electroplating individual metal and alloy dots as the microjet rasters over a substrate. The rastering of a microjet just a few microns above the substrate provides extraordinary convective mass transport rates, allowing material growth rates that are routinely 2 orders of magnitude greater than in conventional plating. EcP has the flexibility to vary the print resolution and material composition during the patterning process via real-time control of the microjet fly height, electrolyte flow rate, and applied current. Microjet fly height has been used, for example, to vary the print resolution from 50 to 1000 dots per inch (dpi) during the printing of a Cu pattern. Simultaneous control of electrolyte flow rate and applied current through the microjet nozzle was used to achieve high plating efficiencies, well-formed dots, and alloys of specific compositions for the Cu and Ni–Cu systems.

Recently, L. Hirt et al. (2016) demonstrated the combination of the EcP process and an atomic force microscope (AFM). This allowed fabrication of 3D metal structures with in situ control of the growth by the AFM control loop. A special silicon AFM probe (FluidFM®, see https://www.cytosurge.com/page/homepage) is filled with a metal salt and immersed in an electrochemical cell. The probe is connected to a high-precision pressure regulator that controls the solution flow rate from the probe orifice. The tip is positioned at a distance of 500 nm from the substrate, and when a sufficiently cathodic potential is applied – metal ions exiting the tip are reduced locally. When the metal deposit reaches the probe, it is recognized by the deflection of the cantilever, the deposition stops, and the tip changes its location. In this manner, no calibration is required since the growth is automatically detected. The printing is carried out layer by layer. Several 3D copper objects were fabricated, having a width of several microns, aspect ratio of up to 1 : 75, and angles as high as 90° (see Figure 19.4).

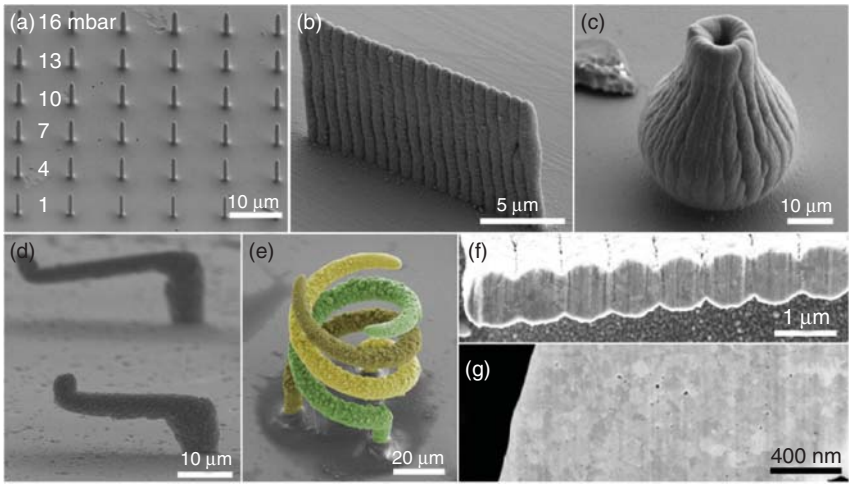


Figure 19.4 SEM images of various copper shapes manufactured by EcP. (a) Array of pillars fabricated at various overpressure values (approximate fabrication time: 20 minutes). (b) Wall structure of 1 µm thickness and 15 µm height (15 minutes). (c) A hollow vase (80 minutes). (d) Arms at 90° and 100° angles, demonstrating overhanging structures. (e) A triple helix, demonstrating the capability to fabricate intertwined shapes thanks to the layer-by-layer approach. The individual helices are colored for better distinction (30 minutes). (f) In-lens detector SEM image of a focused ion beam (FIB) cut cross-section of the wall in (b). (g) FIB cross-section through a Cu pillar. Source: Reprinted with permission from L. Hirt et al., *Adv. Mater*. **28** (2016) 2311.

Dip-pen nanolithography (DPN) is an AFM tip-based process that is used for transferring "ink" molecules directly from the tip on to a surface. Ink molecules are transported via mass diffusion from the tip on to the surface. Features are formed on the surface via either chemisorption-based self-assembly or physisorption. Ink transport is mediated by the presence of the water meniscus that forms when the tip is brought in close proximity to the surface and water molecules in the air are condensed. DPN was first introduced by Chad Mirkin et al. in 1999. Since then, it has been used to pattern a wide variety of materials, from small organic molecules to DNA and metal ions. In *electrochemical dip-pen nanolithography (E-DPN)*, the deposition is limited to surface patterning. In spite of its high resolution, E-DPN suffers from some significant limitations such as very slow deposition rate, limited amount of material on the tip, and no real 3D capabilities.

The *meniscus-confined electrodeposition (MCED)* process, first introduced by T. Leïchlé et al. (2006), is an electrodeposition method that uses the thermodynamic stability of a liquid meniscus to directly print 3D microstructures. This process is also known as *electrochemical fountain pen nanofabrication (ec-FPN)*. MCED consists of long-travel piezo stages that enable very precise control of movement of a micropipette containing an electrolyte solution along the desired 3D trajectory. Dispensing micronozzles with internal diameters ranging from hundred nanometers to tens of microns can be mounted onto the micropipette in order to control the feature size of the structures. The micropipette is moved

toward the conductive substrate, and an electrical potential is applied between a small wire inside the pipette and the substrate. At the appropriate distance, the meniscus is formed between the substrate and the micronozzle; thus, electrodeposition is initiated onto the substrate. The dispensing micronozzle is then moved away from the substrate at a calibrated speed that matches the metal deposition speed in order to keep meniscus formation between the nozzle and the deposited materials, allowing continuous fabrication. Several properties affect the final quality and size of the features: the mechanical stability of the physical system (wettability), the nozzle's diameter, its moving speed, the thermodynamic properties of the electrolyte solution, the interaction between the electrodeposit and the substrate surface, and the applied overpotential. MCED is a flexible, low-cost process (compared to traditional lithography techniques) that is capable of fabricating 3D structures both in nano- and micro-scales, including freeform helical microstructures (down to few microns) under room conditions. However, materials suitable for this technique are limited to those that can be electrochemically deposited with the use of an electrolyte solution (Cu, Pt, Co, Ni, and Au have been successfully used). In addition, it requires high calibration of the parameters to form meniscus. Potential applications include high-frequency terahertz antenna (high aspect ratio > 100), precision sensors, high-density interconnects for integrated circuits, high-aspect-ratio AFM probes for critical metrology, nanoscale needle probes or probe arrays, etc.

The wettability of both the substrate and the tip are very important to the MCED process. Fang et al. (2006) found that when using a hydrophilic tip, the surface energy of the substrate has a great impact on the deposition: a hydrophilic substrate causes the entire reservoir to instantaneously be transferred, while as the substrate is more hydrophobic the deposit droplet decreases in diameter. When modifying the surface energy of the tip, by different chemical depositions, a control of the deposit diameter is achieved. As the tip's outer surface becomes more hydrophobic, the meniscus created on the tip is highly confined, and the resultant deposit size decreases. It was also shown that, when using a hydrophilic tip, the droplet's diameter is closely related to the diameter of the orifice, and is about twice that of the tip opening. While the tip approach/withdrawal velocity and the contact force were found to be insignificant, the contact time changed the amount of liquid transferred to the substrate. The droplet size increased until the contact angles on the tip and substrate reached equilibrium with the hydrostatic pressure imposed by the reservoir. This equilibrium was attained faster and with smaller final volumes as the tip became more hydrophobic.

A relation between nozzle diameter and deposition rate (in the z-direction) was reported by S. Morsali et al. (2017). The deposition rate was found to decrease with nozzle diameter. For example, a stable deposition was maintained with 0.52–2.3 μm s^{-1} nozzle speed for a 100 nm pipette, whereas 15–45 nm s^{-1} was found appropriate for a 5 μm pipette. In addition, the smallest and largest deposited wire diameters possible were claimed to be 0.5 D_0 and 0.9 D_0, respectively, where D_0 is the nozzle diameter. A relation between meniscus height and deposited wire diameter as well as a region for stable deposition were defined, as shown in Figure 19.5.

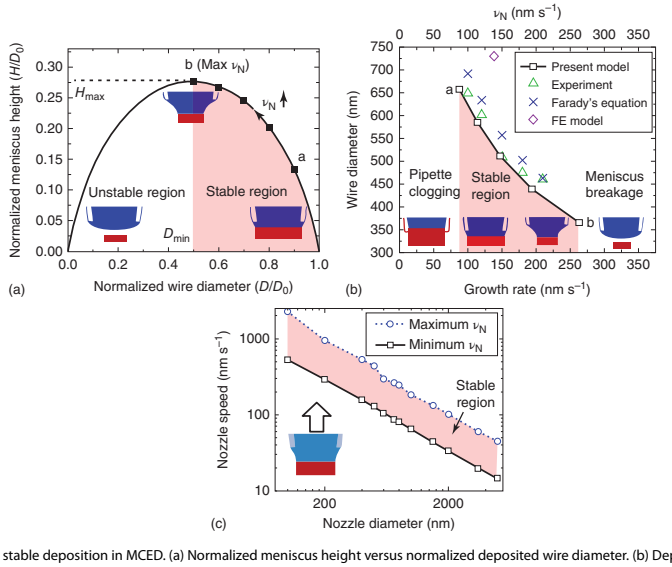


Figure 19.5 Regions for stable deposition in MCED. (a) Normalized meniscus height versus normalized deposited wire diameter. (b) Dependence of the diameter of the deposited wire on nozzle speed ($D_0 = 730$ nm). (c) Minimum and maximum permissible values for nozzle speed. Source: Reprinted with permission from S. Morsali et al., *J. Appl. Phys*. **121** (2017) 214305.

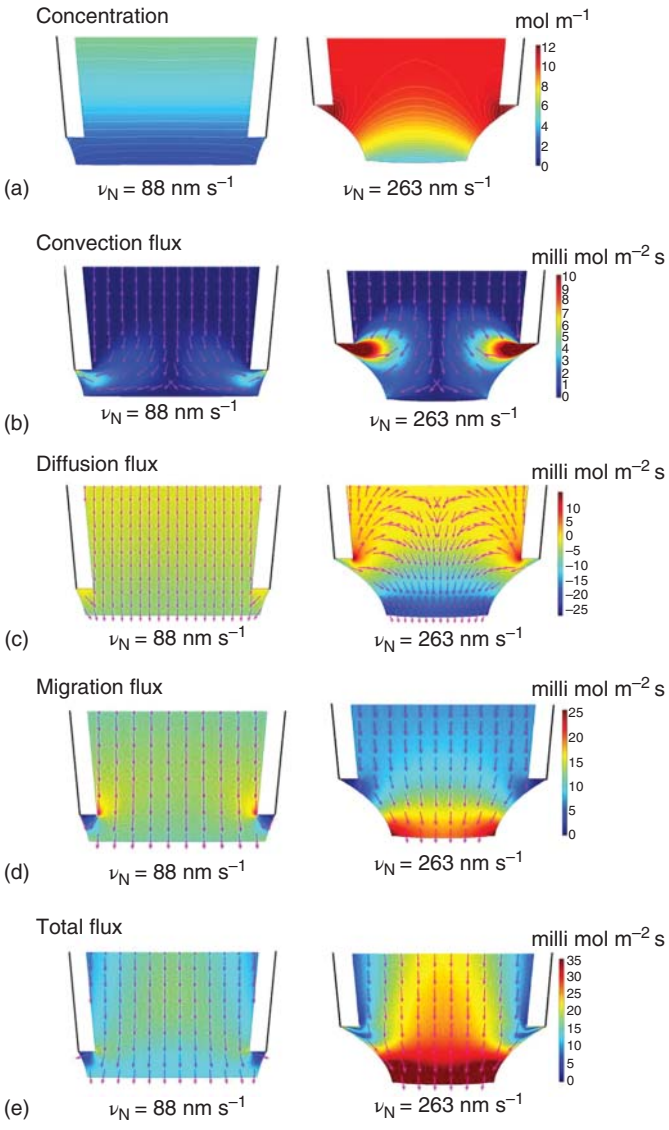


Figure 19.6 (a) Cu^{2+} ion concentration maps around the meniscus at two different nozzle speeds (88 and 263 nm s^{-1}). Maps of different components of the ions flux close to the wire's growth: (b) convection flux, (c) diffusion flux, (d) migration flux, and (e) total flux. The diameter of the nozzle is 730 nm in this simulation. Source: Reprinted with permission from S. Morsali et al., *J. Appl. Phys*. **121** (2017) 214305.

Maps of ion concentration and flux components are presented in Figure 19.6 (for Cu^{2+} ions and nozzle speeds of 88 and 263 nm s^{-1}). As expected, the concentration of Cu^{2+} is higher at the faster nozzle speed (Figure 19.6a). Higher convection flux is generated on the meniscus top corner because the evaporation rate is higher in that region (Figure 19.6b). Convection carries the Cu^{2+} ions to

the meniscus surface and generates a diverging flux that increases the concentration at the vicinity of the meniscus free surface (see Figure 19.6a). Diffusion carries the ions from the electrolyte bulk and meniscus surface toward the wire's growth front due to concentration gradient, and facilitates a more uniform concentration distribution, which results in a converging ion flux (Figure 19.6c). At 263 nm s^{-1}, however, back diffusion happens since the concentration of Cu^{2+} ions in the meniscus top corner is higher than the bulk solution (Figure 19.6a). At 88 nm s^{-1}, all the ions brought by convection and migration are consumed by electrodeposition, because the electrodeposition consumes more ions since the cross-section of the wire (cathode surface) is 1.8 times larger than that at 263 nm s^{-1}. Hence, the concentration of Cu^{2+} within the meniscus remains considerably less than in the bulk electrolyte, which triggers ions diffusion from the electrolyte bulk solution to the wire's growth front. Migration and diffusion are the dominant flux transport mechanisms near the wire's growth front (see Figure 19.6c,d). The direction of total flux (Figure 19.6e) is from the bulk electrolyte to the wire's growth front and has its highest value close to the centerline of the meniscus, in particular near the wire's growth front, due to the contribution of migration and diffusion.

J. Hu and M.F. Yu (2010) applied MCED to deposit Pt columns with diameter as small as 80 nm. They also developed an automated wire bonding process. The resistivity of the copper bridges was measured as ca. 3×10^{-8} Ω m, their breakdown current density higher than 10^{11} A m^{-2}, and their strength higher than ~39 MPa.

Z. Yi et al. (2016) have demonstrated the capability of MCED to print switches made of copper wires (see Figure 19.7). The minimum distance of 500 nm between the wires' upper part is limited by the pipette geometry. The energy dispersive spectrometer (EDS) maps (Figure 19.7c) show that although they were fabricated in ambient environment, only little amount of oxidation occurred.

19.3.2 Nanoelectrochemistry Using Micro- and Nano-Electrodes/Pipettes

There are different ways in which nanometer-scale dimensions can be introduced into electrochemical systems. First, it is possible to scale down the size of the working electrode such that its lateral dimensions are nanometric. These experiments are essentially equivalent to those performed with UMEs, except that the smaller dimensions of the electrodes yield further enhancements in performance, for example, the ability to achieve a true steady state in extremely short times, as well as the possibility of reaching higher current densities. The latter is particularly relevant for measuring fast heterogeneous kinetics. Second, it is possible to establish conditions in which two or more electrodes or surfaces (each of which may or may not be nanometric in size) are separated by nanometer-scale distances. This has been most famously exploited in thin-layer cells and in the so-called positive feedback mode of scanning electrochemical microscopy (SECM) , in which oxidation and reduction taking place at separate, closely spaced electrodes lead to enhanced mass transport (P.S. Singh et al., 2011).

The small size of a nanoelectrode also enables to probe faradaic reactions in extremely tight spaces, such as inside a single vesicle, a biological cell, or a single

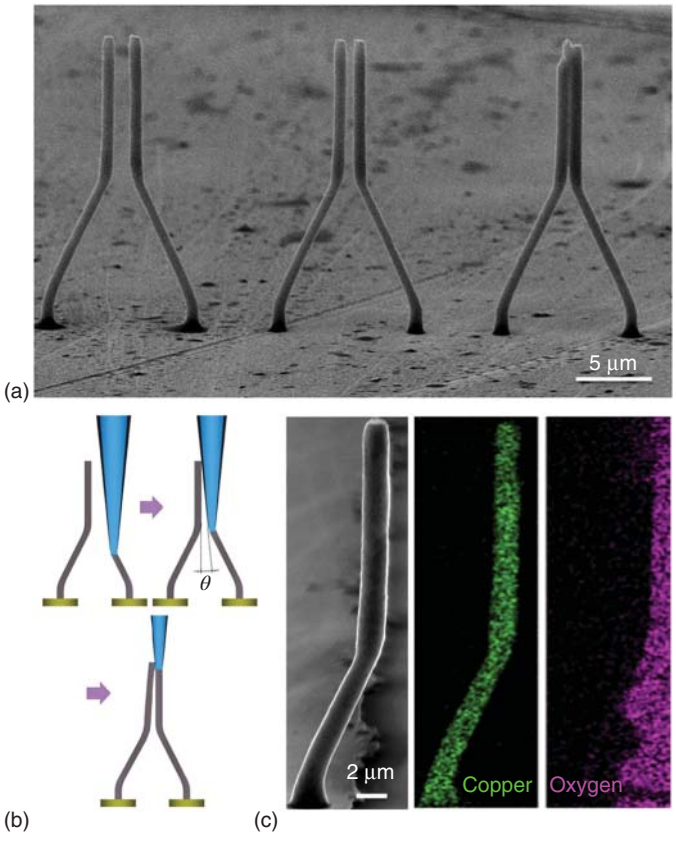


Figure 19.7 Growth evolutions of 3D microswitches. (a) Micrographs of switches with different active gap distances, ranging from 2 to 0.5 µm (left to right). (b) Schematics of the failure mode due to liquid-bridge formation during the fabrication process. (c) Energy dispersive spectrometer (EDS) maps of a single microwire. Source: Reprinted with permission from Z. Yi et al., *Microsys. Nanoeng.* **2** (2016) 16010.

droplet, which are otherwise difficult or even impossible to measure with conventional electrodes. The miniaturized electrode size also leads to insignificant electrical double layer capacitance and a small *RC* time constant. For example, a 10 nm electrode may have a double-layer capacitance of ~30 aF and an *RC* time constant of the order of 30 fs in an electrolyte solution such as 0.1 M KCl. With such a small electrode, one can expect to run ultrafast voltammetric measurements using scan rates in the megavolt per second range and probe electrochemical processes in the nanosecond timescale. Another attractive feature of a nanoelectrode is the minimized *jR* drop due to the small faradaic current, often of the order of picoamperes (pA). This allows one to perform electrochemistry in a solution with very high electrical resistance containing little or even no added supporting electrolyte (Y. Fan et al., 2016).

A number of effects have to be considered for the quantitative analysis of electrochemical phenomena at the nanoscale. Most of these effects are derived

from the observation that during a heterogeneous reaction, the thickness of the diffusion layer decreases linearly with decreasing electrode size. In nanometric electrodes, the diffusion layer and the electrochemical double layer cannot be treated separately from each other, as in classical theory. As the electrode dimensions decrease, the electrochemical double layer occupies an increasingly large fraction of the diffusion layer, and electroactive species start to interact with the interfacial electric field. In the presence of an excess of supporting electrolyte, when the Debye length is short, this effect can be neglected for electrodes larger than 10 nm. However, with smaller electrodes or in the absence of supporting electrolyte, the mass transport rates may be increased or decreased by the electric field within the electrochemical double layer, depending on the charge of the electroactive species with respect to the electrode surface.

Pipettes can be categorized by their volume (e.g. microliter or nanoliter) or by the tip diameter, which is more appropriate for electrochemical purposes. According to the IUPAC, electrodes with a diameter of 0.2–20 μm are referred to as microelectrodes, while those with smaller diameters are referred to as nanoelectrodes. Owing to their simplicity of fabrication, inherent stability, and unique thermal properties, borosilicate and quartz are the two most common materials for manufacturing of micro- and nanopipettes, the latter often being preferred due to its higher toughness and lower electrical noise. The most common way of pumping fluids in microchannels is either by using a syringe pump (constant volume) or a pneumatic pump (constant pressure). Presently, the state-of-the-art syringe pumps can reach picoliters per minute flow rates ($1\ \text{pL} = 10^{-12}\ \text{L}$). Precision small-angle stepper motors are used to reach such low values. Despite such capability, they suffer from oscillation in the flow due to stepping action and slow response at low flow rates. In pneumatic pumps, the pressure inside the fluidic channel is controlled instead of the flow rate. A pulse-free flow and fast response time (tens of milliseconds) can be achieved with this technique. The flow rate has to be calibrated for applied pressure for every given channel dimensions.

The small dimensions of nanopipettes affect their electrochemical behavior. When using glass nanopipettes filled with a low-concentration electrolyte, higher currents are measured at negative potentials compared to the corresponding positive potentials. This effect is termed ion current rectification (ICR), and is attributed to the formation of a diffuse electrical double layer within the nanopipette tip orifice (C. Wei, A.J. Bard, and S.W. Feldberg, Anal. Chem. **69** (1997) 4627). When the double-layer thickness is comparable to the nanopipette diameter, the electrostatic interaction between an ionic species and surface charges will affect ion transport properties. The electrostatic interaction has been shown to be influenced by electrolyte concentration, pH, and applied voltage, and can be modulated by functional layers deposited or covalently attached to the nanopipette orifice. ICR is only observed at electrolyte concentrations at or below 100 mM; no significant current rectification has been reported at high electrolyte concentrations.

Small electrolyte concentrations might also introduce noise to the current measurement. As the number of ions in the solution gets lower, the effect of concentration fluctuations as a result of deposition and diffusion to and from the reservoir will be more noticeable, introducing an intrinsic noise. Another

special consideration is attributed to the small currents flowing through the nanopipette. Because such small currents will not polarize a counter electrode, a two-electrode configuration is more common, considering the counter electrode as a pseudo-reference electrode. Since nanobubbles are a major source of electrical noise in solid-state nanopores, the electrolyte solution is usually degassed prior to use. The noise can be further reduced by treating glass pipettes with sulfuric acid/hydrogen peroxide ("piranha solution") before laser pulling.

The size of the droplets dispensed from the micro- and nanopipettes is determined by a variety of factors, mainly liquid properties (Reynolds number, surface tension, capillary forces, viscosity, Laplace pressure, concentration); tip–surface interactions (Bond number, capillary length, adhesion mechanism, molecular transport mechanism, contact angle, contact time, number of contacts on the same spot, tip's withdrawal speed); environment conditions (humidity, temperature, evaporation rate); and geometrical properties (tip shape, nozzle size, nozzle position). If the nanopipette is withdrawn from the substrate at higher velocities than those required for a continuous and smooth deposition, the meniscus can break and the deposited shape might either stop growing or grow defected (e.g. with beaded parts). The evaporation of electrolyte near the tip of the pipette, which is exposed to the ambient environment, tends to form crystallites on the tip, which might clog the nanopipette, thus preventing further deposition. To prevent this crystallization, the humidity of the adjacent environment should be properly controlled. The use of low-concentration electrolyte can also mitigate the problem of clogging, but will result in reduction of the ion current and the deposition rate. The applied potential should be more negative than the reduction potential of the deposited metal, but preferably less negative than the hydrogen evolution line associated with hydrolysis of water, because significant hydrogen bubble formation can agitate the meniscus and prevent stable deposition.

In fluid mechanics, capillary length, λ_c, is a characteristic length scale for an interface between two fluids (i.e. liquids and gases), which is subject both to gravitational acceleration and to a surface force due to the surface tension of the fluid–fluid interface, σ. It is defined as

$$\lambda_c = \sqrt{\frac{\sigma}{\rho g}} \tag{19.11}$$

where g is the standard acceleration of gravity and ρ is the density of the fluid. When the characteristic height of the liquid meniscus in a liquid bridge between the surface and the tip is quite smaller than the capillary length, the effect of the hydrostatic pressure due to gravity can be neglected. This is also expressed in terms of the Bond number, B_o (also called the Eötvös number, E_o), a dimensionless number used in fluid dynamics to express the importance of gravitational forces compared to surface tension forces, as reflected by the shape of bubbles or drops moving in a surrounding fluid:

$$B_o = \frac{\Delta \rho g L^2}{\sigma} \tag{19.12}$$

where L is the characteristic length of a capillary surface (height or radius of a droplet, or radius of a channel) and $\Delta\rho$ is the difference in density of the two phases. When $B_o < 1$, surface tension forces dominate, and gravitational forces may be neglected. For a dispensed droplet (sessile drop) with a spherical cap shape, the largest dimension is the base of the droplet, which is smaller than the capillary length. The capillary length of clean water at standard temperature and pressure is around 2.7 mm, whereas the channel dimensions of an AFM cantilever are in the micrometer range. Therefore, the effect of gravity on the fluid flow inside hollow channels may be neglected in this case.

The contact angle (wettability) is one of the crucial parameters when dispensing in a nonliquid, non-mixable environment. It is sensitive to a monolayer of molecules on the substrate, and is extremely difficult to handle and control. The Young's equation is the basis of all models for determining the surface free energy of solids by means of contact angle measurements:

$$\gamma_{SG} = \gamma_{SL} + \sigma_{LG} \cos\theta \tag{19.13}$$

where θ is the contact angle of the droplet with the substrate, σ_{LG} the surface tension of the liquid, γ_{SL} the interfacial tension between solid and liquid, and γ_{SG} the surface free energy of the solid. These three forces acting at the triple line per length give the droplet its shape. This equation is valid for three-phase systems in thermodynamic equilibrium for ideal (smooth and chemically homogenous) solids and pure liquids.

In a sufficiently narrow (i.e. low Bond number) tube of circular cross-section (radius r), the interface between two fluids (liquids and gases) forms a meniscus that is a portion of the surface of a sphere with radius R. The pressure jump across this surface is

$$\Delta p = 2\sigma R \tag{19.14}$$

where Δp is the pressure difference across the fluid interface. This may be shown by writing the Young–Laplace equation

$$\Delta p = \sigma \left(\frac{1}{R_1} + \frac{1}{R_2} \right) \tag{19.15}$$

in spherical form with a contact angle boundary condition and also a prescribed height boundary condition at, say, the bottom of the meniscus. Here, R_1 and R_2 are the radii of curvature of the meniscus in orthogonal directions. In order to maintain hydrostatic equilibrium, the induced capillary pressure is balanced by a change in height, h, which can be positive or negative, depending on whether the wetting angle is less than or greater than 90°.

If a loaded pipette (tip) is brought in contact with a hydrophilic substrate, the liquid will start spreading out. Retracting the pipette quickly before reaching equilibrium between the substrate's surface energy and the surface tension of the fluid inside the meniscus will stop such spreading. Furthermore, if the nanopipette possesses a hydrophilic outer wall, the liquid being spread on the pipette wall will increase the dispensing spot diameter much beyond the orifice size, because the meniscus extends from the wall of the nanopipette rather than from its orifice only. To prevent the wetting of the outer tip wall on hydrophilic

substrate surfaces, cantilever tips deposited with gold are often chemically treated (e.g. with hexadecanethiol) to form a hydrophobic outer tip wall. This allows the deposition of smaller droplets.

Further Reading

1 Seh, Z.W., Kibsgaard, J., Dickens, C.F. et al. (2017). Combining theory and experiment in electrocatalysis: insights into materials design. *Science* 355: eaad4998.
2 Morsali, S., Daryadel, S., Zhou, Z. et al. (2017). Multi-physics simulation of metal printing at micro/nanoscale using meniscus-confined electrodeposition: effect of nozzle speed and diameter. *J. Appl. Phys.* 121: 214305.
3 Morsali, S., Daryadel, S., Zhou, Z. et al. (2017). Multi-physics simulation of metal printing at micro/nanoscale using meniscus-confined electrodeposition: effect of environmental humidity. *J. Appl. Phys.* 121: 024903.
4 Hirt, L., Reiser, A., Spolenak, R., and Zambelli, T. (2017). Additive manufacturing of metal structures at the micrometer scale. *Adv. Mater.* 29: 1604211.
5 Fan, Y., Han, C., and Zhang, B. (2016). Recent advances in the development and application of nanoelectrodes. *Analyst* 141: 5474–5487.
6 Hirt, L., Ihle, S., Pan, Z. et al. (2016). Template-free 3D microprinting of metals using a force-controlled nanopipette for layer-by-layer electrodeposition. *Adv. Mater.* 28: 2311–2315.
7 Yi, Z., Guo, J., Chen, Y. et al. (2016). Vertical, capacitive microelectromechanical switches produced via direct writing of copper wires. *Microsys. Nanoeng.* 2: 16010.
8 Mirkin, M.V. and Amemiya, S. (2015). *Nanoelectrochemistry*. Boca Raton, FL: CRC Press.
9 Mathwig, K., Aartsma, T.J., Canters, G.W., and Lemay, S.G. (2014). Nanoscale methods for single-molecule electrochemistry. *Annu. Rev. Anal. Chem.* 7: 383–404.
10 Farahani, R.D., Chizari, K., and Therriault, D. (2014). Three-dimensional printing of freeform helical microstructures: a review. *Nanoscale* 6: 10470–10485.
11 Alkire, R.C., Kolb, D.M., and Lipkowski, J. (ed.) (2014). *Electrocatalysis: Theoretical Foundations and Model Experiments*. Wiley.
12 Eliaz, N. (2011). *Applications of Electrochemistry and Nanotechnology in Biology and Medicine I & II*, Modern Aspects of Electrochemistry, vol. 52 and 53. New York: Springer.
13 Osaka, T., Datta, M., and Shacham-Diamand, Y. (ed.) (2010). *Electrochemical Nanotechnologies*. New York: Springer-Verlag.
14 Hu, J. and Yu, M.F. (2010). Meniscus-confined three-dimensional electrodeposition for direct writing of wire bonds. *Science* 329: 313–316.

15 Schmuki, P. and Virtanen, S. (ed.) (2009). *Electrochemistry at the Nanoscale*. New York: Springer-Verlag.
16 Lin, Y. and Nalwa, H.S. (ed.) (2007). *Handbook of Electrochemical Nanotechnology*. American Scientific Publishers.
17 Piner, R.D., Zhu, J., Xu, F. et al. (1999). "Dip-pen" nanolithography. *Science* 283: 661–663.

20

Energy Conversion and Storage

20.1 Introduction

Energy production and consumption that rely on the combustion of fossil fuels is forecast to have a severe future impact on global economics and ecology. Therefore, renewable energy sources, such as solar energy, wind energy, and hydropower, are being developed. At the same time, the development of high-energy storage systems is essential to save extra power for the increased demands and delivery to where it is required. There are some basic factors that are vital to energy storage systems, such as the following:

1. *Capacity* (A h): The amount of energy that the cell can deliver in a single discharge. It is calculated by multiplying the discharge current by the discharge time.
2. Cost.
3. *Cycle life*: The number of discharge–charge cycles the cell can experience before it fails to meet specific performance criteria – normally, 80% of the rated capacity.
4. *Energy density* (W h L^{-1}): The nominal cell energy per unit volume, also called the volumetric energy storage density.
5. *Open-circuit voltage* (V): The voltage between the cell terminals with no load applied.
6. *Power density* (W L^{-1}): The maximum available power per unit volume, also known as the volumetric power density of the cell.
7. Safety.
8. *Specific energy* (W h kg^{-1}): The gravimetric energy storage density of the cell. It is a characteristic of the battery chemistry and packaging.
9. *Specific power* (W kg^{-1}): The maximum available power per unit mass, where the power generated by the cell is the multiplication of cell current by cell voltage.

Energy storage technologies available for large-scale applications can be categorized as mechanical, electrical, chemical, and electrochemical. The latter has become very popular because it possesses a number of desirable features, including pollution-free operation, high roundtrip efficiency, a long life cycle, low maintenance, and energy characteristics such as quick response when the contingency

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

occurs to meet different grid functions. Systems for electrochemical energy storage and conversion include batteries, fuel cells (FCs), and supercapacitors – also known as electrochemical capacitors (ECs). Although the energy storage and conversion mechanisms are different in these systems, they do have certain "electrochemical" similarities. Common features are that the energy-providing processes take place at the phase boundary of the electrode/electrolyte interface, and that electron and ion transport are separated. In this chapter we will discuss batteries, fuel cells, and supercapacitors, as well as hydrogen storage.

20.2 Batteries

20.2.1 Classes of Batteries

In batteries and fuel cells, electrical energy is generated by conversion of chemical energy via redox reactions at the anode and cathode. The difference between batteries and fuel cells is related to the locations of energy storage and conversion. Batteries are closed systems in which the anode and cathode undergo the redox reaction; thus, energy storage and conversion occur in the same compartment. Fuel cells are open systems where the anode and cathode are just charge-transfer media and the active masses undergoing the redox reaction are delivered from outside the cell, either from the environment or from a tank (oxygen from air and fuels such as hydrogen are to respective examples). Hence, in fuel cells, energy storage (in the tank) and energy conversion (in the fuel cell) are locally separated.

All batteries are composed of two electrodes connected by an ionically conductive material – the electrolyte. The two electrodes have different chemical potentials, dictated by the chemistry that occurs at each. When these electrodes are connected by means of an external device, electrons spontaneously flow from the more negative to the more positive potential. Ions are transported through the electrolyte, maintaining the charge balance, and electrical energy can be tapped by the external circuit. The amount of electrical energy per mass or volume that a battery can deliver is a function of the cell's voltage and capacity, which are dependent on the chemistry of the system. Another important parameter is power, which depends partly on the battery's engineering, but crucially on the chemicals the battery contains.

The stored energy content of a battery can be maximized in three ways: (i) by having a large chemical potential difference between the two electrodes; (ii) by making the mass (or volume) of the reactants per exchanged electron as small as possible; and (iii) by ensuring that the electrolyte is not consumed in the chemistry of the battery. For example, one of the key elements of Ni–MH and Li-ion batteries is that the same ion (H^+ for the former, Li^+ for the latter) participates at both electrodes, being reversibly inserted and extracted from the electrode material, with the simultaneous addition or removal of electrons.

We may thus consider three classes of batteries:

(a) *Primary batteries*, which are designed for use only once. The chemical energy is stored in the two electrodes, or sometimes in the electrolyte,

and it is converted to electrical energy on demand. Until about the middle of the twentieth century, the most important use of primary batteries was for flashlights. With the introduction of portable electronic devices, laptops, remote controls, initially radios, tape recorders and cameras, toys, and nowadays just about everything from civilian to military applications, primary batteries are manufactured in immense quantities worldwide. In 2014, 93.5% of the battery units sold in the United States were primary. In terms of global market share, however, they account for less than 34%.

This is not necessarily an efficient way of storing energy. The total energy consumed in the manufacturing of a primary battery may be several times as much as the energy that can be retrieved from it. Yet, primary batteries offer by far the best way to store energy in small packages, and in many applications they constitute the only way to store energy.[1] The first battery was the *Leclanché cell*, invented by Georges Leclanché in 1866. Nowadays, the so-called *alkaline cell* – a variant of the Leclanché cell – is by far the most common primary battery. In both cases, the chemical energy is stored in the form of metallic Zn at the anode and MnO_2 at the cathode, as will be discussed below. In the second half of the twentieth century, primary batteries based on Li anodes and a nonaqueous solvent were introduced. Such batteries are better than the Leclanché cell in almost every aspect other than cost.

(b) *Secondary (rechargeable) batteries*, as their name indicates, are designed for multiple use. A larger voltage applied in the opposite direction can cause the battery to recharge. The *lead–acid battery*, used in our cars as well as in many other applications, has the largest market (in USD) among all secondary batteries. Another rechargeable battery in use for many decades is the *Ni–Cd battery* consisting of metallic Cd as the anode and trivalent nickel oxide as the cathode, in concentrated KOH solution. The batteries used in modern laptops and other portable electronic devices are based on a *rechargeable Li-ion battery* (LIB) and a nonaqueous solvent.

(c) A *fuel cell* is commonly regarded as a different type of energy conversion device, but it is just another type of battery. For example, the difference between a lead–acid battery and a hydrogen–oxygen fuel cell is in the way the energy is stored and replenished. As explained before, in the battery, it is all internal: the chemical energy is stored in the two electrodes, and the battery is recharged, converting electrical energy to chemical energy. In the fuel cell, the chemical energy is stored in containers outside the electrochemical cell. Recharging is replaced by the addition of chemical energy (hydrogen and oxygen in this case) from outside containers. When the device is being used as a source of electric power, chemical energy is being converted to electrical energy in both cases.

The unique feature of batteries and fuel cells is that electrical energy is produced in them *directly* from chemical energy, bypassing the need to convert it first to heat, and then construct a heat engine of one type or another to convert the thermal energy to electrical energy. The efficiency of conversion of the latter

1 It has been noted, rather sarcastically, that if all batteries would be destroyed by a magic wand, war as we know it could not be waged.

is limited by the second law of thermodynamic, expressed by the Carnot cycle,[2] as given by

$$\eta = \frac{T_2 - T_1}{T_2} \tag{20.1}$$

In the above equation, T_2 and T_1 represent the temperatures of the hot and the cold heat reservoir in the system. It follows that thermal energy cannot be extracted (or converted to any other form of energy) isothermally. In contrast, a fuel cell converts chemical energy directly to electrical energy isothermally, according to the simple equation

$$\Delta G = -nFE \tag{17.38}$$

This equation shows that, at the limit of reversibility, the Gibbs energy released by the system can be converted to electrical energy with 100% efficiency. A similar process, of converting chemical energy to mechanical energy, takes place in the human body and most other living organisms. The energy needed to move our muscles is derived from the chemical energy of oxidation of food by the oxygen we breathe.

The ability to oxidize (burn) fuel isothermally and at high efficiency, not limited by the Carnot cycle, is very attractive. Although this principle was already demonstrated in the middle of the nineteenth century, serious efforts to develop fuel cells started only in 1959, when a group of scientists claimed that propane gas could be used directly to produce electric power in a fuel cell. About a decade later, the first successful implementation of a H_2—O_2 fuel cell was made in the Apollo space program, aimed at sending men to the Moon. Since then, great progress has been made in the development of fuel cells. At present, the most common system being studied is a *direct methanol fuel cell (DMFC)*. Fuel cells will be described more thoroughly in Section 20.3.

20.2.2 The Theoretical Limit of Energy per Unit Weight

Thermodynamics allows us to calculate the maximum specific energy of a battery, which may be approached, but never reached, in real batteries. This type of calculation is based on the reversible potential and on the equivalent weight of the electrochemically active ingredients, ignoring the current collector, the casing, etc. The standard potential has to be substituted into the Nernst equation, which takes into account the nonstandard state of the reacting component. The Nernstian potential will change with time either because of use or self-discharge by which the activity of the electroactive component in the cell is modified. Thus, the nominal voltage is determined by the cell chemistry at any given point of time. The actual voltage will always be lower than the theoretical voltage due to polarization and resistance losses (IR drop) of the battery, and is dependent upon the load current and the internal impedance of the cell. These factors are

2 Both the internal combustion engine and the major electrical power station are "heat engines" in the thermodynamic sense, and their theoretical maximum efficiency is that of the Carnot cycle.

dependent upon electrode kinetics, and thus vary with temperature, state of charge, and with the age of the cell.

As an example, consider the Ni–Cd rechargeable battery, which we will discuss in more detail in section titled "The Nickel–Cadmium Battery". The overall reaction in this battery is

$$2\mathrm{NiO(OH)} + \mathrm{Cd} + 2\mathrm{H_2O} \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} 2\mathrm{Ni(OH)_2} + \mathrm{Cd(OH)_2} \tag{20.2}$$

The standard cell potential is 1.299 V. The electrical energy produced is given by $nFE = 2 \times 96{,}485 \times 1.299 \approx 250.7\ \mathrm{kJ\ mol^{-1}} \approx 69.6\ \mathrm{W\ h\ mol^{-1}}$.[3] The sum of the molecular masses of the reactants is 331.8 g $\mathrm{mol^{-1}}$, leading to a *theoretical* specific energy of $69.6/0.3318 = 209.8\ \mathrm{W\ h\ kg^{-1}}$. This should be compared to the actual value of 40–60 W h $\mathrm{kg^{-1}}$, namely, 19–29%.

What is the purpose of this kind of calculation? In the development of a battery it is important to know how far we are from its theoretical limit, since the practical limit, which is typically half the theoretical limit, can be reached asymptotically, and the effort in approaching it grows exponentially. Thus, if current technology represents only 20% of the theoretical limit, there is a very good probability that it could be developed to 50%, namely, by a factor of 2.5. If, on the other hand, current technology has reached 45% of the theoretical limit, it is probably close to its practical limit, and the wisdom of attempting to develop it further may be questioned.

20.2.3 How is the Quality of a Battery Defined?

There are many parameters by which the quality of a given type of battery can be judged. Most important are the specific energy in units of W h $\mathrm{kg^{-1}}$ and the power density in units of W $\mathrm{kg^{-1}}$. But, in certain applications volume is more important than weight. This applies, for example, to batteries for wrist watches and for hearing aids. For possible application in electric cars, weight is the more important parameter, but volume could also play an important role.

For primary batteries, the rate of self-discharge that determines shelf life is critical, while for rechargeable batteries a higher rate of self-discharge may be tolerated. Safety for the user is always a major issue. Imagine the battery in your laptop catching fire (in your lap, of course, where else!), particularly if this might happen on a commercial flight. Reliability is particularly critical for some applications, such as heart pacemakers or the communication system of an army unit in battle. Price is more important in civilian than in military and space applications, but is sometimes less critical for *in vivo* medical applications. Finally, environmental issues during manufacturing, use, and disposal can no longer be ignored.

It should be remembered that primary batteries, born more than a century ago in a world obsessed by the development of new technologies, may die soon in a world obsessed by ecology and recycling. Indeed, it is almost unbelievable that in the first decade of the twenty-first century we still use and throw away many

3 Recall that 1 W h = 3600 J.

billions of batteries every year, while the technology to produce rechargeable batteries, which could be reused thousand times, already exists. Considered from the point of view of recycling, the transition from primary to secondary batteries is equivalent to recycling paper, aluminum, or beer cans thousands of times.

20.2.4 Primary Batteries

20.2.4.1 Why Do We Need Primary Batteries?

It cannot be overemphasized that batteries (of any kind) are not sources of energy – they are devices to store energy. The energy needed to build a primary battery far exceeds the energy that can be stored in it. Moreover, primary batteries are extremely expensive when the cost is calculated per unit of energy. For example, the cost of AA-size alkaline cells needed to store 1 kW h may be anywhere between $100 and $1000, compared to about $0.15 paid to the electric utility company for the same amount of energy. On the other hand, primary batteries carry energy in "small change" and do that more efficiently than any other device. Since the battery in your quartz-crystal watch lasts for two or three years and costs a few dollars, who cares about the cost per kW h?

20.2.4.2 The Leclanché and the Alkaline Batteries

The oldest commercial primary battery is the Leclanché cell. In its modern version it consists of a Zn container acting as the anode (the negative electrode), where metallic Zn is oxidized to $Zn(OH)_2$. The electrolyte is NH_4Cl or $ZnCl_2$, with a small amount of water to allow electrolytic conductivity, yet leaving it in the form of a paste or gel. The open-circuit potential (OCP) is about 1.5 V. The cathode material is MnO_2, mixed with carbon powder and in intimate contact with a carbon rod acting as the current collector, and is the positive electrode in the cell. The reactions taking place in a Leclanché cell are

$$Zn + 2H_2O \rightarrow Zn(OH)_2 + 2H^+ + 2e^-_M \tag{20.3}$$

$$2MnO_2 + 2H^+ + 2e^-_M \rightarrow 2MnOOH \tag{20.4}$$

resulting in the overall reaction

$$Zn + 2MnO_2 + 2H_2O \rightarrow Zn(OH)_2 + 2MnOOH \tag{20.5}$$

Zinc is an active metal ($E^0 = -0.763$ V versus standard hydrogen electrode (SHE)) that corrodes in aqueous solutions, giving off molecular hydrogen according to the reaction

$$Zn + 2H_2O \rightarrow Zn(OH)_2 + H_2 \tag{20.6}$$

This reaction, which takes place at open circuit or as a side reaction during discharge of the battery, is detrimental in two ways: (i) it consumes one of the active materials in the cell, and (ii) it produces a gas that could build up pressure and eventually rupture the cell. This is the purpose of adding HgO to the zinc anode. In contact with metallic zinc, it is reduced to metallic mercury, which amalgamates the zinc. The exchange current density for hydrogen evolution is much lower on mercury and its amalgam than on zinc, thus reducing the rate of

self-discharge via the hydrogen evolution reaction (HER). Other inhibitors have also been used, but mercury is the most effective in this respect. In early designs, the mercury content of the anode in the final product reached a good percentage. This amount has been gradually reduced, and eventually mercury was eliminated altogether because of its toxicity.

Leclanché cells have several disadvantages. They have a relatively short shelf life, and must be refrigerated for long-time storage. The specific energy is about 75 W h kg^{-1}, which is relatively low for a primary battery, and the power density is also low. The voltage at constant load declines during discharge, which is a disadvantage for most applications, although it affords a convenient way to monitor the state of charge of the battery. On the other hand, the greatest advantage of Leclanché cells is their low price. They are therefore widely used for simple applications where reliability and performance are not of critical importance.[4]

A similar system, containing the same active materials, is the so-called *alkaline battery*. This type of battery differs from the Leclanché cell in that the electrolyte is concentrated KOH, zinc is in the form of a high-surface-area paste, and the casing is made of nickel-coated steel. It has a much higher power density, because of the higher conductivity of the electrolyte and the larger surface area of the zinc anode, and is less likely to leak. However, it is more expensive. Alkaline batteries are the most widely sold primary batteries nowadays.

20.2.4.3 The Li–Thionyl Chloride Battery

For about a century after the invention of the Leclanché cell (and the lead–acid rechargeable battery), the highest voltage attained in any battery was 2.1 V, and primary cells delivered a maximum voltage of about 1.5 V. This changed in the mid-1960s of the twentieth century, when the Li–thionyl chloride ($Li/SOCl_2$) battery was introduced. In this and other nonaqueous lithium batteries, a nominal voltage of 3.6 V is currently common.

The $Li/SOCl_2$ battery consists of a lithium anode and a carbon-paste cathode. The electrolyte consists of a solution of $LiAlCl_4$ in $SOCl_2$. The anode reaction is metal dissolution:

$$Li^0_M \rightarrow Li^+_{soln} + e^-_M \tag{20.7}$$

The cathode reaction is more complicated, since in this battery there is no reducible material in the cathode itself, and it is the solvent that serves as the active cathode material:

$$2\,SOCl_2 + 4\,e^-_M \rightarrow SO_2 + S + 4Cl^- \tag{20.8}$$

and the overall cell reactions is

$$4Li^0 + 2\,SOCl_2 \rightarrow SO_2 + S + 4LiCl \tag{20.9}$$

The reversible cell voltage was estimated to be 3.65 V. Assuming the above cell reaction, this leads to a theoretical specific energy of 1.48×10^3 W h kg^{-1}. Values as high as 700 W h kg^{-1} have been realized in commercial cells. Unlike the

4 In Leclanché cells the casing is made of zinc and serves as the anode. This design increases to some extent the chances of leakage, but makes manufacturing very cheap.

Leclanché cell, Li–thionyl chloride batteries are designed to be anode limited, i.e. to have a stoichiometric deficiency of lithium, since discarding a spent cell with a residue of metallic Li could create a serious safety hazard!

Similar batteries having SO_2 or propylene carbonate (PC) as the solvent have been prepared. The former have better characteristics at low temperature. In the case of PC, the solvent does not take part in the reaction, and MnO_2 is used as the cathode material. This yields a slightly lower voltage of 2.8 V, but PC is much easier to handle than either $SOCl_2$ or SO_2, both in the manufacturing process and for waste disposal.

Lithium *primary* batteries have many advantages compared to Leclanché cells and other aqueous batteries, but they are also interesting from the fundamental point of view. To begin with, the Gibbs energy of interaction of lithium with the solvent is so high that they would be expected to react violently with each other, leading to a very high rate of self-discharge at best, and to dangerous explosions at the worst. However, as soon as contact between the metal and the solvent is made, a protective layer is formed, which prevents further chemical reaction. In the $Li/SOCl_2$ and Li/SO_2 cells, this layer consists of LiCl, which is not soluble in the solvent. If PC is used as the solvent, the protective layer consists of Li_2CO_3. Fortunately, these layers are permeable to Li^+ ions, but not to electrons. In fact, the protective layer on the surface of Li serves as an electrolyte, having a transference number of unity with respect to the positive ion. It is referred to in the literature as the solid/electrolyte interphase (SEI).

A layer formed on a metal upon contact with the solution could be of three different types. If it is dense and nonconducting, it can protect the metal from corrosion, but the system cannot be used as a battery, since the metal is totally isolated from the solution. If it is electronically and ionically conducting, reduction of the solvent at the film/electrolyte interface and oxidation of the metal at the metal/film interface could proceed freely, leading to a high rate of self-discharge. It is only when the film is simultaneously an *ionic conductor* and an *electronic insulator* that the chemical pathway of spontaneous reduction of the solvent at the anode is blocked, whereas the electrochemical pathway of oxidation of the metal at the anode and reduction of the solvent at the cathode can proceed at a sufficient rate to allow the use of the system as an anode in a battery.

In addition to the high specific energy of $Li/SOCl_2$, the operating cell voltage in low-rate batteries is around 3.6 V, more than twice that of aqueous primary batteries. As a result, computer memories based on complementary metal–oxide–semiconductor (CMOS) technology can be operated using a single cell.

The rate of self-discharge of primary Li batteries is very low, about 10 times lower than that of alkaline cells, allowing storage without refrigeration for periods of 5–10 years. Another welcome feature is the stability of the voltage during discharge, which is shown in Figure 20.1a. This is an obvious advantage for the operation of any electronic device. The only drawback is that the voltage at open circuit or under load cannot be used as a measure of the state of charge of the battery, and special devices have to be developed to determine the state of charge of $Li/SOCl_2$ and other lithium-based nonaqueous batteries.

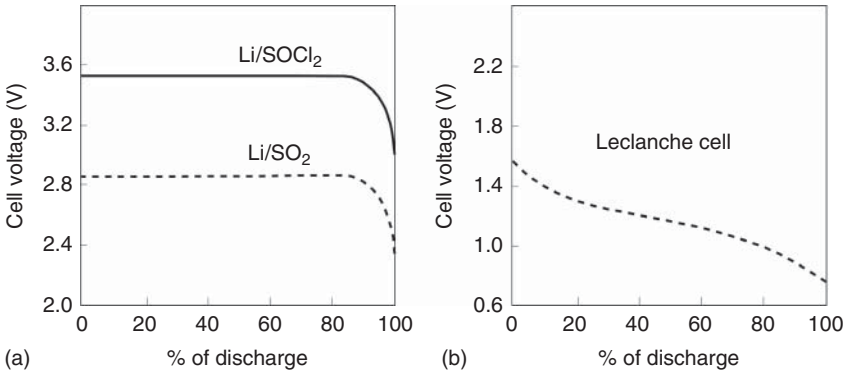


Figure 20.1 The discharge curves of (a) $Li/SOCl_2$ and $LiSO_2$ cells, and (b) Leclanché cell. Note the stability of the voltage during most of the discharge process for the Li cells, compared to the gradual decrease of voltage of the Leclanché cell. Source: Reprinted with permission from D. Linden, *Handbook of Batteries and Fuel Cells*, D. Linden (ed.), McGraw-Hill, Chapter 3, 1984.

Ideally, one would like to have a flat discharge curve, namely, a potential that is almost constant throughout the discharge stage and falls fairly sharply when the battery has been exhausted. This is an obvious advantage for the operation of any electronic device using such a battery. The Li cell is obviously better than the Leclanché cell, as seen in in Figure 20.1. However, the decline of the voltage of the Leclanché cell makes it easy to measure the rate of declension of power, as seen in Figure 20.1b, while in the case of the Li battery, the voltage is almost steady until about 80% discharge, and measuring the remaining performance is difficult until it is too late, as seen in Figure 20.1a.

The high specific energy of this type of $Li/SOCl_2$ battery combined with the low rate of self-discharge and the high operating voltage makes them ideally suited to power the memory backup in computers. For this usage, the battery is designed to outlast the useful lifetime of the computer itself, and probably will never have to be replaced.

For low-power applications, these batteries are quite safe, but high-power lithium batteries have been known to explode when accidental heating melts the lithium ($T_m = 180.5\,^\circ C$). This can rupture the protective SEI layer, leading to a violent reaction between the metal and the solvent, and eventually to explosion.

20.2.4.4 The Lithium–Iodine Solid-State Battery

This type of battery is used for heart pacemakers. The electrode reactions in this case are very simple, leading to the overall cell reaction

$$2\,Li + I_2 \rightarrow 2\,LiI \tag{20.10}$$

The electrolyte is solid LiI, which is formed *in situ* during operation of the cell. The solid electrolyte interface is conducting for Li^+ ions, but nonconducting for electrons. The cathode consists of a compound such as poly-2-vinylpyridine, mixed with molecular iodine and melted together until a homogeneous material has been formed. This constitutes a charge-transfer complex, allowing easy transfer of the iodine. The thickness of the solid LiI layer grows during discharge of the

battery, and its resistivity is quite high, but not too high, considering the very low currents (in the range of 10 μA or less) needed to operate the pacemaker. The cell voltage is 2.8 V,[5] and the specific energy is about 250 W h kg^{-1}. The advantages of this type of battery are its very low rate of self-discharge and its extremely high reliability. Although it has been used in tens of thousands of cardiac patients over several decades, there has not been a single case of failure causing lethal internal injury, to the best of our knowledge. For the very special requirement of heart pacemakers, the foregoing advantages make this type of battery commercially viable, in spite of its very high cost.[6]

20.2.5 Secondary Batteries

20.2.5.1 Self-Discharge and Specific Energy

Secondary batteries are rechargeable. The rate of self-discharge is less critical in this case, because the battery can always be recharged before use. Cycle life, which is the number of times the battery can be charged and discharged, is important. The efficiency of energy storage in secondary batteries is quite high, on the order of 60–90%. A good battery can be charged and discharged a thousand times with little loss in performance. Specific energy, power density, and the temperature range over which the battery can be operated are important. The specific energy of a battery depends, of course, on the rate at which it is discharged, and it must be defined for a specific rate. Three factors are involved here: (i) The activation overpotential depends on the current density during discharge (i.e. on the rate of discharge). (ii) Some potential is lost as a result of the resistance of the solution and of the separator in the cell. (iii) The amount of available charge also depends on the discharge rate.

20.2.5.2 Battery Stacks Versus Single Cells

A battery stack consists of a number of cells connected in series and packaged as a unit. This is required when the voltage of a single cell is not high enough to operate the device. It is invariably observed that the reliability in performance and the lifetime of stacks of batteries is less than that of single cells. This can be understood on the basis of very simple statistical reasoning. If the probability of failure of a single cell during the intended service life of a stack is 1%, the probability for its flawless operation is 0.99. The probability of trouble-free operation of a stack consisting of 12 cells of this type connected in series is $0.99^{12} = 0.89$. Thus, a reliability of 99% for the individual cell translates to a lower reliability of only 89% for the stack! Repeating the same calculation for a battery stack containing 100 cells in series yields a reliability of only $0.99^{100} = 0.366$. On the other hand, if the probability of failure is reduced to 0.1% for a single cell, a stack of 100 cells will have a reliability of $0.999^{100} = 0.904$.

---

5 According to tables of standard potentials, the OCP of this cell should be 3.58 V. Note, however, that standard potentials are given for aqueous solutions, whereas here the reaction occurs in the solid state, where the Gibbs energy of the reaction can be quite different.

6 Most of the cost is due to the extremely high reliability demanded in this kind of application. For other uses, such as memory backup in computers, the same battery could be manufactured at a much lower cost.

20.2.5.3 Some Common Types of Secondary Batteries

The Lead–Acid Battery The best known secondary battery with the largest market is the lead–acid battery, which was invented in 1859 by Gaston Planté. It consists of a metallic Pb anode and a PbO_2 cathode, both immersed in a solution of about 25% sulfuric acid. Upon discharge, the anode is oxidized to $PbSO_4$, and the cathode is reduced to the very same compound. The reaction at the anode during discharge is

$$Pb + SO_4^{2-} \rightarrow PbSO_4 + 2e_M^- \tag{20.11}$$

while at the cathode lead dioxide is being reduced, yielding the same product

$$PbO_2 + 2\,H_2SO_4 + 2\,e_M^- \rightarrow PbSO_4 + SO_4^{2-} + 2H_2O \tag{20.12}$$

Thus, the overall cell reaction is given by

$$PbO_2 + Pb + 2\,H_2SO_4 \underset{\text{charge}}{\overset{\text{discharge}}{\rightleftharpoons}} 2\,PbSO_4 + 2\,H_2O \tag{20.13}$$

We note that sulfuric acid is being consumed during discharge, and an equivalent amount of water is formed. The resulting decrease in density of the electrolyte used to be the basis for the method used in many garages to test the state of charge of a car battery by measuring this density.

The open-circuit cell voltage is 2.1 V, the highest voltage for any *aqueous* battery. In fact, we would not expect a battery having such a high voltage to hold charge at all, because the voltage is high enough to electrolyze water, providing an efficient way for rapid self-discharge. Indeed, if the terminals of a lead–acid battery are connected to two platinum electrodes placed in a cell containing the same solution, copious evolution of hydrogen and oxygen is observed. But, lead is not platinum. The kinetics of hydrogen evolution on lead is slow, leading to a very high overpotential for this reaction. Thus, the operation of the lead–acid battery depends, to a large extent, on the sluggishness of the HER at the cathode. This is a case in which having a bad electrocatalyst is an asset rather than a liability! The successful operation of the lead–acid battery depends, of course, on the rate of the overall reaction shown in Eq. (20.13) in both directions.

The theoretical specific energy of lead–acid batteries is only 171 W h kg^{-1}, due to the high atomic mass of lead. The practical specific energy depends on the rate of discharge, but even at low rates it does not exceed about 40 W h kg^{-1}. This represents about 23% of the theoretical value, despite massive investments in engineering aimed at increasing the specific energy of this type of batteries.

As pointed out above, the rate of self-discharge of secondary batteries is not as important as it is for primary batteries. One does not expect a car battery to be fully charged after the car has been idle for several months. The quality of a secondary battery is measured, instead, in terms of its service life, which is determined by the number of times it can be charged and discharged. To understand this limitation, one must consider the structure of the electrodes in the cell and the way in which they are charged and discharged. As a rule, the active material is in the form of a high-surface-area powder, pressed onto the surface of a metallic

grid, which serves as the current collector. The anode in the lead–acid battery consists of small particles of lead bonded to a lead screen.[7] During extended use, and in particular as a result of abuse, the particles may agglomerate, causing loss of surface area. Now, we note that during discharge lead is converted into $PbSO_4$, which is an insulator. If a lead particle is coated with its oxidation product, it becomes isolated from the solution and can no longer take part in the discharge process. Agglomeration increases the size of the particles, causing loss of active surface area and, thus, decreasing the effective charge capacity of the battery. The same mechanism can cause deterioration of the cathode as well, since PbO_2 is electronically conducting, while the $PbSO_4$ formed on the surface of these particles during discharge is not.

Another common mode of failure of batteries is loss of electrical contact between the active material and the current collector. There are many other ways in which batteries can fail, such as the aging of separators and accidental contact between anode and cathode. These problems are not discussed here.

The largest asset of lead–acid batteries is their low cost, compared to any other secondary battery currently available. The specific energy is inherently low, but in its current application as a car battery, this is tolerable. For application as the main power source of electric vehicles (EVs), batteries with higher performance are being developed now (2018). It can be stated that an electric family car would not have a lead–acid battery as its main source of energy. Power density is another limitation, particularly since increasing the power can decrease the specific energy dramatically.

The lead–acid battery reigned supreme among rechargeable batteries for about a century, because it is constructed from inexpensive materials and is highly reliable. When properly treated, a car battery can last for several years. Proper treatment in this context means that it is not allowed to dry out; it is never discharged completely and never overcharged. It can also deliver a large power (of 1.0–1.5 kW) for a few seconds, as needed to start a car on a very cold day. The efficiency of storage and retrieval of energy, which is often called the electric-to-electric (ETE) efficiency, can be as high as 80%, depending on the way the battery is used.

Lead–acid batteries are very heavy (in terms of energy stored per unit weight), but this is of little consequence when a single battery is used in a car. On the other hand, it makes it essentially impossible to use such batteries as the power source for modern electric cars, although there are some niche applications, such as golf carts and vehicles for indoor operations, such as in airport terminals or for fork lifts inside factories.

The Nickel–Cadmium Battery The Ni–Cd battery was invented by Waldemar Jungner in 1899. It consists of a sponge Cd anode and a NiO(OH) cathode, in a

7 The real situation is invariably more complex. The screen is not made of pure Pb, but may contain some Ca or Sb. A binding material is used to increase adhesion of the lead particles to the screen. Each manufacturer has his own secret formulations, but we do no need to know these to understand how such batteries operate.

concentrated solution of KOH.[8] The cathodic and anodic reactions taking place during discharge of this cell are, respectively,

$$NiO(OH) + H_2O + e_M^- \rightarrow Ni(OH)_2 + (OH)^- \tag{20.14}$$

$$Cd + 2(OH)^- \rightarrow Cd(OH)_2 + 2e_M^- \tag{20.15}$$

The standard reduction potentials of these two half cells are +0.490 and −0.809 V, respectively. The exact composition of the fully charged cathode is not entirely clear. Some prefer to write it as NiO_x, where $1.5 \leq x \leq 2$, implying that there may be some Ni^{4+} mixed with Ni^{3+}. There seems to be little doubt that the cathode can be charged beyond the level corresponding to the conversion of all the active material to trivalent nickel, but the existence of tetravalent nickel has not been confirmed. The overall reaction is

$$2NiO(OH) + Cd + 2H_2O \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} 2Ni(OH)_2 + Cd(OH)_2 \tag{20.2}$$

The nominal operating cell voltage is 1.2 V, the cell's standard potential is 1.299 V, and the theoretical specific energy is 208 W h kg^{-1}. Discharge curves for a Ni–Cd battery are shown in Figure 20.2. This figure shows clearly the two effects that reduce the energy obtained from a battery as a function of the rate of discharge. On the one hand, the cell voltage is decreased with increasing discharge rate. On the other hand, the fraction of the active material that can be utilized is decreased. For a 0.1 C (10-hour discharge), the potential is almost the same as for the reference rate of 1 C, but the total charge available is about 10% higher. For a rate of 10 C (6-minute discharge) the output voltage is decreased to about 1.0 V, and the available charge is decreased by about a factor of 2.

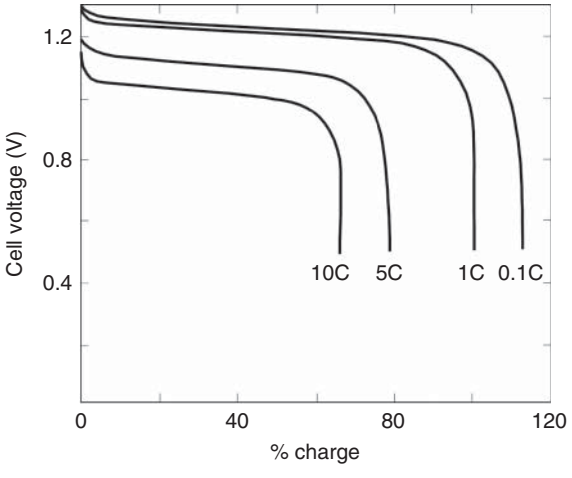


Figure 20.2 Typical discharge curves of a Ni–Cd battery at different rates. Source: Reprinted with permission from J.M. Evjen and A.J. Catotti, in *Handbook of Batteries and Fuel Cells*, D. Linden (ed.), McGraw-Hill, Chapter 17, 1984.

8 This again is an oversimplification. The cadmium anode may contain a small percentage each of iron, nickel, and graphite. The NiO(OH) cathode may contain some Co and also graphite. The electrolyte is 20–34% KOH, but some LiOH may be added to increase the conductivity of the electrolyte.

Here, we draw attention to the special way the discharge rate is defined for batteries in the literature. Thus, a discharge rate of 5 C means that the cell is discharged in 1/5 hours, while a rate of 0.1 C indicates that it is discharged in 1/0.1 hours. Full discharge is defined as the charge measured at 1 C. This is somewhat arbitrary, but nevertheless quite useful as a point of reference.

The main advantage of Ni–Cd over lead–acid batteries is the longer cycle life. At an 8-hour discharge rate (0.125 C), the two batteries may be nearly equal in specific energy, but at a 30-minute rate (2 C) the Ni–Cd battery still performs well whereas the lead–acid battery can barely work, losing 80% of its capacity. The performance of high-quality Ni–Cd batteries as a function of the number of cycles is shown in Figure 20.3. About 10% of the capacity is lost in the first thousand cycles, but no further decline in capacity is observed up to about 2500 cycles. Performance deteriorates rapidly beyond 3500 cycles, and this can be considered to be the limit of the useful cycle life of such batteries. In terms of real time, Ni–Cd batteries can last 5–10 years, depending on the design, quality, and type of application. In comparison, lead–acid batteries typically do not last more than about 500 cycles or five years. Another advantage of the Ni–Cd battery is that its weight per unit energy stored is lower.

In defining the cycle life of different batteries, one should be careful to state how the test was performed. Charging back and forth from 10% to 90% may be a reasonable choice, but for a particular case that test may not be relevant. For example, a lead–acid battery in a car may be charged most of the time with occasional heavy use (when the car is started), so that its cycles could be from 95% to 75%. In comparison, a Ni–Cd cell in a camera may be used typically until it is almost totally discharged, representing cycles of 5% to 95%. The implication of this is that in contrast to the lead–acid battery, it is necessary to discharge the Ni–Cd battery completely from time to time, in order to maintain good performance. The worst drawback of the Ni–Cd battery is its cost, but progress has been made in recent years, and Ni–Cd rechargeable batteries are gradually replacing

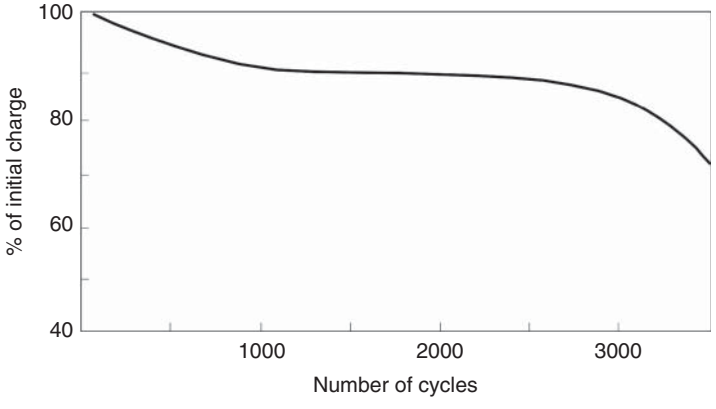


Figure 20.3 The decline in charge capacity with cycle life for a high-quality sealed Ni–Cd battery. The cycle used was five hours charge and three hours discharge at 0.4 A. Discharge was stopped when the potential reached 1.0 V. Source: Data from J.A. Wiseman, in *Handbook of Batteries and Fuel Cells*, D. Linden (ed.), McGraw-Hill, Chapter 18, 1984.

Leclanché cells and alkaline Zn/MnO_2 primary batteries in many simple applications, such as mechanical toys and flashlights.

The Nickel–Metal Hydride Battery (NiMH) Nickel–metal hydride (NiMH) batteries are essentially extension of the sealed Ni–Cd battery technology with the substitution of a hydrogen-absorbing negative electrode for the Cd-based electrode. While this substitution increases the battery's electrical capacity (measured in A h) for a given weight and volume and eliminates Cd with its toxicity issues, the remainder of this battery is quite similar to the Ni–Cd battery. An important advantage of NiMH batteries lies upon their cycle life; typically, they can be recharged hundreds of times, potentially allowing them to be equivalent to hundreds of alkaline batteries in total service over their lifetime. However, battery life is limited to two to five years. The specific energy of the NiMH battery is in the range of 40–110 W h kg^{-1}, depending on the power requirements, which can be as high as 1900 W kg^{-1}. This battery has proved to be a great success, with sales already exceeding one billion dollars a year, and it is gradually replacing Ni–Cd batteries. The rechargeable NiMH batteries are a cost-effective power source for many frequently used battery-operated devices found at home or in the office.

There is a long way to go between storage of dry hydrogen and building a working rechargeable battery with $LaNi_5$ as the anode and the medium for storing hydrogen. Indeed, the first attempts to build such a battery with the NiO(OH) cathode were not promising. It was only the persistence and perseverance of Ovshinsky and the extensive research conducted in his company (Ovononics) that led to the development of practical NiMH rechargeable batteries that could compete successfully with the Ni–Cd battery.

The anode of a current NiMH battery on the market has the typical AB_5 structure, but both A and B do not represent pure metal, but rather some nonstoichiometric combination of metals of the same group as La and Ni, respectively. A typical formulation can be written as[9] $[La_{5.7}Ce_{8.0}Pr_{0.8}Nd_{2.3}]$ $[Ni_{59.2}Co_{12.2}Mn_{6.8}Al_{5.0}]$, where the numbers represent the relative atom percent concentration of each element. Adding up the numbers, this corresponds to $A_{16.8}B_{83.2}$, which could also be written as $AB_{4.95}$, very close to the $LaNi_5$ used above.

The unique feature of this complex formulation is that the relative concentrations of elements in groups A and B can be varied in order to optimize the performance of the battery, aiming at maximum energy, highest power, longest cycle life, or lowest cost. The electrolyte is alkaline, between 20% and 40% (by weight) solution of KOH, containing other minor constituents to enhance battery performance. Typically, the separator, which provides electrical isolation between the electrodes while still allowing efficient ionic diffusion, is a nonwoven polyolefin.

The electrochemistry of the NiMH battery can be represented as follows. During charging, water in the electrolyte is decomposed at the cathode (negative electrode) into hydrogen atoms, which are absorbed in the alloy, and hydroxyl anions:

$$M + H_2O + e_M^- \rightarrow MH + (OH)^- \tag{20.16}$$

9 See M.A. Fetcenko, S.R. Ovshinsky, et al., *J. Power Sources*, **165** (2007) 544–551.

At the positive electrode (anode), oxidation of nickel hydroxide occurs just as in the Ni–Cd battery:

$$Ni(OH)_2 + (OH)^- \rightarrow NiO(OH) + H_2O + e_M^- \tag{20.17}$$

During discharge, hydrogen is desorbed at the anode (negative electrode) and combines with a hydroxyl ion to form water:

$$MH + (OH)^- \rightarrow M + H_2O + e_M^- \tag{20.18}$$

At the cathode (positive electrode), nickel oxyhydroxide is reduced to its lower valence state, nickel hydroxide:

$$NiO(OH) + H_2O + e_M^- \rightarrow Ni(OH)_2 + (OH)^- \tag{20.19}$$

20.2.5.4 The Li-ion Battery

The introduction of the Li/thionyl chloride primary battery represented a major advancement in battery technology. The voltage of a single cell was more than twice that of the Leclanché cell and other primary batteries, specific energy was increased by an even higher factor, and shelf life was improved by a factor of 5, at least. It became immediately obvious that the next breakthrough would be achieved with the introduction of a *rechargeable lithium battery*. This happened about two decades later, with the introduction of the LIB, which now has the fastest growing market among batteries.

The *LIB* was invented in 1985 by Akita Yoshino, and was first commercialized by Sony in 1991. The LIB consists of a cathode, an anode, an electrolyte separator, and an outer casing (see Figure 20.4). The exact mass and chemical composition of each component vary between different manufacturers. The *cathode* comprises 25–30% of the battery's total weight and consists of a thin aluminum current collector sheet and a layered metal oxide of the form $Li_{1-x}TMO_2$, where TM is a transition metal (Co, Mn, or Ni). Other examples include $LiMn_2O_4$ and $LiFePO_4$. In the early stages of the development of Li-ion batteries, Co was mainly used, but the higher price and toxicity of Co have led to its common replacement by Mn. The *anode* represents 15–30% of the battery's total weight and is typically made up of a copper current collector sheet that is coated with a graphite layer of the type Li_xC_6. The cathode and anode materials are fused with an inert binder, such as polyvinylidene fluoride (PVF), to help them adhere to their respective metal collector sheets. The separator is usually made of microporous polypropylene or polyethylene, which allows the passage of lithium ions during the cycling process. The separator contains a lithium salt electrolyte (for example, $LiClO_4$, $LiBF_4$, $LiAsF_6$, $LiPF_6$, $Li(CF_3SO_3)$, or $Li[N(CF_3SO_2)_2]$) in an organic solvent typically consisting of a cyclic or linear carbonate, such as ethylene carbonate, PC, or dimethyl carbonate. This solvent typically has a high permittivity and low viscosity. The purpose of the electrolyte is to allow the controlled movement of lithium ions between the electrodes during the cycling process.

When a LIB is fully charged, most of the lithium ions are stored in the anode of the battery. The discharge process is initiated when an external circuit connecting the two electrodes is completed. During discharge, Li is removed from the carbon matrix as Li^+ ion and is intercalated into the cathode, forming compounds

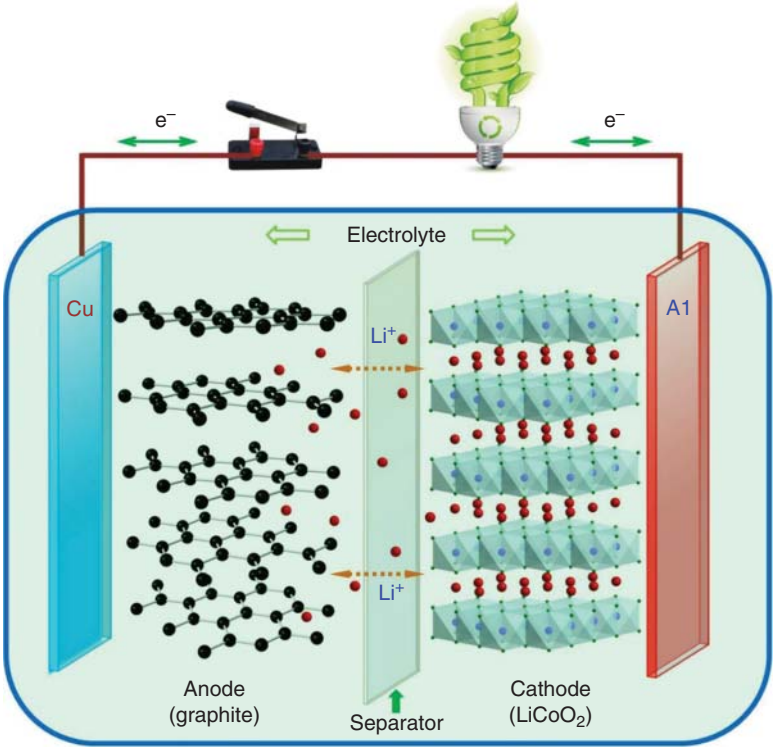


Figure 20.4 Schematic of the configuration of a rechargeable Li-ion battery. Na-ion, Mg-ion, or Al-ion batteries also have similar configurations, but simply differ in electrode materials. For a Li-ion battery, as illustrated in this figure, Li ions are extracted from the cathode and inserted into the anode during the charge process, and the reverse reaction occurs during the discharge process. Source: Reprinted with permission from C. Liu et al., *Materials Today*, **19** (2016) 109.

such as those shown above. The electrolyte separator controls the rate at which the lithium ions can migrate between the two electrodes. While the lithium ions are migrating, electrons are conducted from the anode to the cathode through the external circuit. The electron movement and voltage between the two electrodes is used to power a device. The process is reversible, i.e. when the battery is connected to an external power source, the lithium ions and electrons can be forced to move from the cathode back to the anode.

Contrary to the safety hazards of primary lithium batteries, no metallic Li is used in the rechargeable battery. The processes inside the cell involve the transfer of Li^+ ions from the anode to the cathode during discharge, and vice versa during charging. For this reason, the system was initially referred to as "the rocking chair battery," and it would seem that there was no metallic Li involved either in the manufacturing of the battery or during its operation. Considering the processes taking place at the two electrodes, this is not true. Thus, when the battery is charged, the processes taking place can be written as

$$Li^+_{soln} + 6C + e^-_M \rightarrow LiC_6 \tag{20.20}$$

and

$$LiMn_2O_4 \rightarrow Li^+_{soln} + Mn_2O_4 + e^-_M \tag{20.21}$$

leading to the overall reaction:

$$LiMn_2O_4 + 6C \underset{\text{discharge}}{\overset{\text{charge}}{\rightleftharpoons}} LiC_6 + Mn_2O_4 \tag{20.22}$$

This is an odd reaction for a LIB, considering that the Li^+ ion does not even show up in the overall reaction. But, it does play a role in the two partial reactions at the anode and the cathode.

One may wonder whether there is any metallic Li in this battery, either in the charged or the discharged form. Strictly speaking, the answer is no, but a closer look at Eq. (20.20) shows that it does involve the reduction of a Li^+ ion to metallic Li, except that it is not stored as such, but rather "dissolved" in the graphite. Had we replaced graphite by mercury (which would obviously be a very bad idea, both from the point of view of technology and that of the environment), the same process would lead to an amalgam of Li in Hg, which we would definitely regard as metallic Li. Considering LiC_6 as being similar to an amalgam (and, hence, containing in effect metallic Li dissolved in carbon) is confirmed by measurement of the OCP of the Li-loaded graphite versus a piece of pure Li. This turns out to be about +(0.08 – 0.10) V. Moreover, when a partially loaded LIB is opened and dumped in water, copious hydrogen evolution is observed. Thus, the OCP of a LIB represents the Gibbs energy involved in transferring a Li ion from graphite to Mn_2O_4.

Actually, the cathode in a LIB is generally more complex than indicated above. The process taking place during discharge could be represented by

$$Li^+ + Mn^{iv}_2O_4 \rightarrow LiMn^{iii}Mn^{iv}O_4 \tag{20.23}$$

which shows that intercalation of Li^+ reduces one of the Mn atoms from the four-valent to the three-valent state. In some cases, a fraction of the Mn atoms is replaced by another transition metal, forming nonstoichiometric compounds. Changing the element added and its relative fraction in the oxide can improve the performance of the resulting LIB.

A somewhat different type of LIB has been developed too, employing $LiFePO_4$ as the cathode. In this system, the intercalation of Li^+ into the structure of $FePO_4$ can be represented by reduction of the three-valent iron atom to its two-valent state, as represented by the equation

$$Li^+ + Fe^{iii}PO_4 \rightarrow LiFe^{ii}PO_4 \tag{20.24}$$

The main advantage of this system is in the lower price of Fe compared to Co and even to Mn, and the lower toxicity compared to Co.

This LIB is superior to any other rechargeable system. Its OCP is about 3.6 V, with prospect to increase the voltage up to 4.5 V, compared to 2.1 V for the lead–acid and 1.35 V for the Ni–Cd batteries. Although the specific energy it stores ($\approx$180 W h kg^{-1}) at an average voltage of 3.8 V is only a factor of 5 higher than that stored by the much older lead–acid battery, it required significant R&D in materials science and engineering and in solid-state chemistry to achieve

it. The R&D efforts were aimed at three components of this type of batteries: (i) using different types of graphitic materials for the negative electrode, to increase the amount of Li in the graphite and the rate of its intercalation and de-intercalation; (ii) improving the positive electrode, mostly by replacing a small part of the Mn by other transition metals; and (iii) improving the stability of the solvent, to withstand the higher potential of the cell.

While the specific energy of LIBs is relatively high, so is their cost. Thus, one would not consider using them to replace the alkaline Leclanché cells for flashlights, for example. Instead, they are used where specific energy, reliability, and long service life dominate over cost. LIBs have been used extensively in cell phones, laptops, cameras, and other electronic devices. Nowadays, they are also used in EVs (see Section 20.2.6).

LIBs have some safety issues as they are known to catch fire or explode when overheated, punctured, crushed, mishandled, or when there is a manufacturing defect. The risk of a fire is less of an issue when the batteries are completely discharged because the potential energy within the battery is less than in a charged battery. This has been an issue for garbage and recycling trucks that compact the waste they collect, as this has caused LIBs to explode and cause the waste in the truck to catch fire. Additionally, LIBs have occasionally been reported to cause landfill fires. Therefore, care must be taken when discharging LIBs for recycling purposes.

The capacity of the battery is determined by how many lithium ions can be stored in a given amount of anode material. The theoretical storage capacity of graphite is fairly low, at 372 A h kg^{-1}. A material commonly used as a replacement for graphite is the spinel form of $Li_4Ti_5O_{12}$, which offers a longer cycle life. Carbon nanotubes (CNTs), tin compounds, and metallic nanoparticles (NPs) are among the latest technologies in development for improving anode performance. These new chemistries could increase the recycling value of the anode, which is mostly recycled for the copper current collector.

Attempts to improve the design of Li-ion batteries are nowadays focused on the nanoscale. Nanomaterials were slow to enter the field of energy storage because the effective increase in the electrodes' surface area raised the risk of secondary reactions involving electrolyte decomposition. Only around year 2000 was it realized that such reactions could be controlled by coating the electrodes to protect the electrolyte from unwanted oxidation or reduction by the electrode materials. The introduction of nanomaterials gave Li-ion batteries a new lease of life and provided benefits in terms of capacity, power, cost, and materials sustainability, which are still far from being fully exploited. Nevertheless, the most important downside related to the use of nanomaterials is the poor packing density of electrodes, which limits the energy that can be stored per unit volume or mass, because there is a larger proportion of "inert" components such as current collectors or electrolyte (M. Armand and J.-M. Tarascon, 2008).

A chase is also underway for cathode materials that can undergo conversion reactions involving multiple electrons at high potential. However, the full impact of nanomaterials on living cells has yet to be appraised, although the risk is minimal when they are produced *in situ*, as they are for the conversion reaction.

Solid electrolytes have benefited from the use of nanomaterials. The addition of "nano-fillers" (e.g. Al_2O_3 or TiO_2 nano-grains dispersed in a polymer) to simple polyether-based electrolytes increases the conductivity several fold at 60–80 °C, but there is no advantage at room temperature. Organizing the polymer strands in such a way as to increase the order locally can also provide benefits by increasing conductivity at low temperatures, and further work is needed to assess the merits of using block copolymers (AB or ABA). The phase separation inherent to these systems results in good mechanical properties, but also offers a way of increasing dissociation by partitioning anions and cations in the two subphases. Giving the two phases different wetting or adhesion properties can help by avoiding grain growth as the polymer's nano-domains will determine the partitioning of space (M. Armand and J.-M. Tarascon, 2008).

True polymer batteries may still be some way off, but in the meantime, attempts are being made to use ionic liquids as either solvents for lithium salts or plasticizers for polyether-based electrolytes. Ionic liquids have exceedingly low vapor pressure, are nonflammable, and have high conductivities, making them serious competitors for safer batteries. Yet, it remains to be seen whether they can be produced cheaply enough, at the desired purity, with sufficient conductivity at low temperature (M. Armand and J.-M. Tarascon, 2008).

As evident from Figure 20.5, due to their high energy density and design flexibility, Li-based batteries currently outperform other systems; therefore, they account for more than 60% of worldwide sales values in portable batteries. Note that in Figure 20.5, PLiON stands for "plastic Li ion," the first reliable and practical rechargeable LIB with a thin-film shape and a polymeric electrolyte,

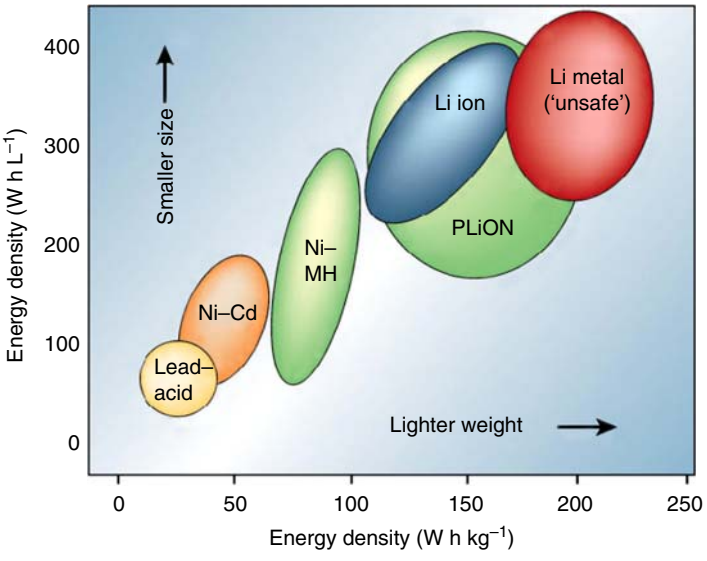


Figure 20.5 Comparison of different battery types in terms of volumetric and gravimetric energy density. It is evident that Li-based batteries currently outperform their competitors. Source: Reprinted with permission from J.-M. Tarascon and M. Armand, *Nature*, **414** (2001) 359.

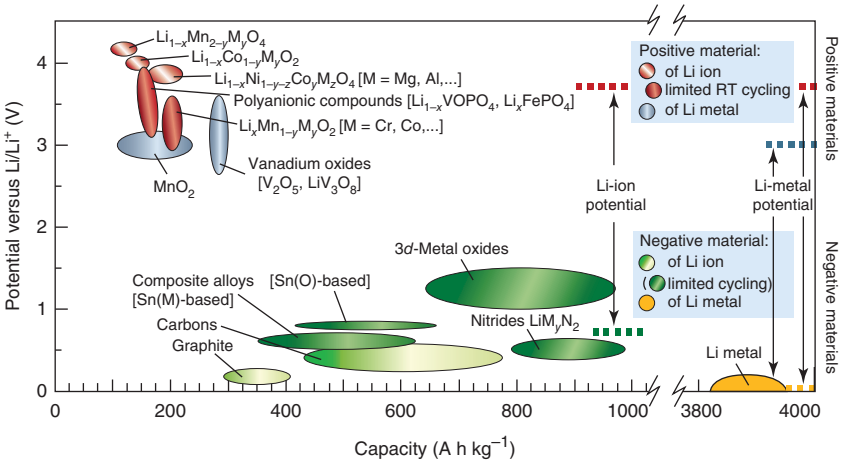


Figure 20.6 Voltage versus capacity for positive- and negative-electrode materials either used or under consideration for the next generation of rechargeable Li-based cells. The output voltage values for Li-ion cells or Li-metal cells are represented. Source: Reprinted with permission from J.-M. Tarascon and M. Armand, *Nature*, **414** (2001) 359.

first introduced by Bellcore in 1999. The membrane in this battery is based on a copolymer of vinylidene difluoride with hexafluoropropylene.

The choice of the positive electrode depends on whether we are dealing with rechargeable Li-metal or Li-ion batteries (Figure 20.6). For rechargeable Li batteries, owing to the use of metallic Li as the negative electrode, the positive electrode does not need to be lithiated before cell assembly. In contrast, for Li-ion batteries, because the carbon negative electrode is empty of Li, the positive one must act as a source of Li, thus requiring use of air-stable Li-based intercalation compounds to facilitate the cell assembly. Although rechargeable Li solid polymer electrolyte (Li-SPE) cells mainly have used Li-free V_2O_5 or its derivatives as the positive electrode, $LiCoO_2$ has been more common in commercial Li-ion batteries, deintercalating and intercalating Li around 4 V. In Figure 20.6, note the huge difference in capacity between Li metal and the other negative electrodes, which is the reason why there has been great interest in solving the problem of dendrite growth.

Improving energy storage and preventing Li_2O_2 from clogging the electrode require a better understanding of the reaction mechanism of the oxygen electrode. Engineering and chemical advances are also required to prevent the ingress of either CO_2 or H_2O, which could react with either Li_2O_2 or lithium metal. But there are reasons for optimism. The use of nanomaterials makes it possible to design porous, catalyzed, 3D electrodes with improved kinetics and energy efficiency. The use of ionic liquids, which can be made hydrophobic, will put an end to problems caused by the entry of water. However, if ionic liquids are to be used as electrolytes, they must be combined with a highly hygroscopic Li salt, so preparing them remains a serious challenge (M. Armand and J.-M. Tarascon, 2008).

The components of today's Li-ion batteries are produced from ores, and not from renewable energy resources. Therefore, extracting the raw materials and

manufacturing the electrodes will require increasing amounts of energy as they become scarcer. Will the LIB, which is so energetically expensive to fabricate, remain attractive and viable in the long run? In 50 years, if all cars become electric and rely on these scarce materials, might we face staggering price increases like those seen with fossil fuels? Not if we find a way of making Li-ion batteries sustainable while maintaining or exceeding the performance of today's batteries. One option is to use renewable electrodes made from natural resources, just as fuel cells can use hydrogen or (m)ethanol made from biomass (M. Armand and J.-M. Tarascon, 2008).

20.2.5.5 Metal–Air Batteries

Metal–air batteries, such as Li–, Zn–, Mg–, and Al–air batteries, are promising for future generations of EVs because they use oxygen gas accessed from the air as their cathode material, thus reducing the weight of the battery and freeing up more space devoted to energy storage. Fundamentally, metal–air batteries can be classified into two types according to their electrolytes: (i) a cell system using an aqueous electrolyte, which is not sensitive to moisture, and (ii) a water-sensitive system using an electrolyte with aprotic solvents, which is sullied by moisture.

Among all the metal–air batteries, Li–air battery shows the highest theoretical energy density, rivaling the gasoline engine. It has shown 5–10 times more energy density than a standard LIB. The estimation of the theoretical capacity and energy density of Li–air batteries, however, is in debate. It was determined that the maximum theoretical gravimetric and volumetric energy densities are 1300 W h kg^{-1} and 1520 W h L^{-1} in basic electrolyte, and 1400 W h kg^{-1} and 1680 W h L^{-1} in acidic electrolyte, respectively. The above specific capacity and energy projections are the theoretical maximum limitation and are based on active materials including only Li metal, air electrode, and electrolyte. However, in practical cells, some excess electrolyte must be in the cell in order to provide ionic conductive media during the entire discharge process. In addition, other necessary materials, including current collectors, membrane, and package materials, will further reduce the cell specific capacity and energy by 20–30%. Furthermore, there are many challenges facing the design of rechargeable Li–air batteries, such as incomplete discharge as porous carbon cathode blocking by discharge products, unstable anode in atmospheric moisture, inadequate understanding of catalysts effect, low electrical efficiency due to higher charge overpotential than discharge overpotential, and carbonate-based electrolytes decomposing during discharge and producing lithium alkyl-carbonates and Li_2CO_3, which severely affects the rechargeability and cycle life of Li–air batteries.

There are also challenges for secondary Zn–air batteries such as requirement of closely controlled zinc precipitation, zinc anode dendrite formation, nonuniform zinc dissolution and limited solubility in electrolytes, higher charge overpotential than discharge overpotential, and necessity of bi-functional air cathode to liberate oxygen from discharge reaction products.

Both Mg– and Al–air batteries are in an early stage of development; while Mg–air batteries are rechargeable, Al–air batteries are not. Although they are attractive having light weight, these two batteries deliver less voltage and confront several challenges such as corrosion problems of metal Mg and Al as

anode, sluggish discharge products, high self-discharge rate, cell irreversibility, and low shelf life. At the same time, only high-capacity cathode materials can be considered, which narrows it down to oxygen or sulfur for Mg, or graphite and fluoride for Al. Yet, little is known about the kinetics of electrode reactions involving the motion of multivalent species, and addressing these challenges would need extensive collaboration between organometal researchers and electrochemists. To conclude, there are still numerous scientific and technical challenges of metal–air batteries that must be overcome before this promising technology is to turn into reality.

20.2.6 Batteries-Driven Electric Vehicles

Since the late twentieth century, environmental pollution and energy crisis have become two global issues, to which the transportation sector is one of the major contributors. Governments have made great efforts, including a series of preferential policies, to support EVs for the benefit of zero emission and energy saving. In 2016, the global EV sales reached 774,000, 40% above that of 2015. The development of energy storage technology, especially LIBs, also greatly accelerated this battery-driven trend in the automobile industry, where they are gradually replacing the NiMH batteries.

At present, the EVs on the market include battery electric vehicles (BEVs), hybrid electric vehicles (HEVs), plug-in hybrid electric vehicles (PHEVs), and extended-range electric vehicles (E-REVs). A BEV is completely powered by an electric motor, and there is no internal-combustion engine (ICE). As a result, BEVs do not produce gas emissions. The requirements from batteries and specifications for BEVs are capacity of 35–54 kW h, specific energy of 90–100 W h kg^{-1} (for LIB), weight of 450 kg (for LIB), vehicle speed higher than 100 km h^{-1}, and a distance on one charge of at least 150–200 km. The batteries of BEVs can mainly be charged with a plug-in charger. Compared to the other three EVs, BEVs have the highest battery capacity and longest pure-electric driving range. In addition, the energy transition efficiency of BEVs is higher than traditional automobiles, and the electricity to charge them can be obtained from wind, the sun, water, and other clean sources of energy. Of course, we first need to convert the above clean energy to electricity, which drives a car. In spite of the longer pure-electric driving range, the total driving range of a BEV is far less than the other three EVs. In addition, due to the longer charging time for BEVs, compared to PHEVs or E-REVs, more care must be given to timely charging in order to avoid a lack of power during the next use. HEVs are each powered by a gasoline ICE and a supplementary electric motor. With its limited battery capacity, the electric motor can only supply power for starting the vehicle and acceleration. HEVs cannot be charged by outlets; instead, the batteries are charged by recovering the energy through braking. The requirements from batteries and specifications for HEVs are capacity of less than 3 kW h, specific energy of 40–50 W h kg^{-1} (for NiMH battery), weight of 60 kg (for NiMH battery), vehicle speed higher than 100 km h^{-1}, and a distance on one charge of 15 km. A PHEV, which is a new type of HEV, has better battery capacity and can be charged from electricity mains. The major difference between PHEVs and HEVs is that the former have

a short pure-electric range, usually about 50–100 km, and can be driven by an ICE after battery depletion. The requirements from batteries and specifications for PHEVs are capacity of 5.6–18 kW h, specific energy of 90–100 W h kg^{-1} (for LIB), weight of 60–200 kg (for LIB), vehicle speed higher than 100 km h^{-1}, and a distance on one charge of 10–60 km. E-REVs, similar to PHEVs, also have better battery capacity and a plug-in charger. When an E-REV depletes its energy in pure-electric mode, a fuel tank is used for extending its driving range by charging the batteries and providing electricity that directly drives the electric motor.

While the highest energy density may be provided by FCs, to be discussed in Section 20.3, it seems that FC technology is not mature enough for practical EV application due to operation problems related to electrocatalysis in direct FCs, as well as some very severe problems in hydrogen storage for H_2/O_2 FCs. Hence, in the visible future, it seems that only Li-ion batteries may have reasonable energy density and cycle life for EV application.

To date, the main barriers to EVs penetration are still low specific energy and specific power, safety issues, cost (related to cycle life and calendar life), low temperature performance (e.g. cold startup in winter), and charge time. Owing to the limited temperature tolerance of components in the Li-ion cells and relatively unstable chemistries, excessive high and low temperatures can both reduce the life and threaten the safety of the battery, to the extent of permanent damage.

EVs already in operation include electric cars, trolleybuses, buses, trucks, bicycles, scooters, tractors, golf carts, and electrified all-terrain vehicles. Already on the last three missions of the Apollo program in 1971 and 1972, astronauts drove silver-oxide battery-powered Lunar Roving Vehicles for distances up to 35.7 km on the lunar surface. Some current specific examples include Toyota Prius (HEV); Nissan Leaf (PHEV); Tesla Model S, Ford Focus Electric, BMW ActiveE, Honda Fit EV, Mercedes-Benz B-Class Electric Drive, and Hyundai Ioniq Electric (BEVs); Chevrolet Volt and Mitsubishi Outlander P-HEV (PHEVs); Tesla Semi (an all-electric battery-powered semi-trailer truck), China's CRH5 high-speed rail, the electric tram in Vienna (Austria), BYD K9 (a battery electric bus with onboard $LiFePO_4$ battery), Solar Impulse 2 (electric, solar-powered aircraft), and many others. As of April 2016, over 11 million HEVs had been sold worldwide since their inception in 1997.

20.2.7 The Polarity of Batteries

The polarity of batteries may sometimes lead to confusion, which we would like to dispel here. The anode and the cathode are defined unequivocally as the electrodes where oxidation and reduction occur, respectively. But, is the anode the positive or the negative terminal of a battery? This question is particularly relevant for the case of secondary batteries, where the electrode serving as the anode during discharge becomes the cathode during charging and vice versa.

Consider the lead–acid battery. The standard reversible potentials for the two half-cells are as follows:

$$PbO_2/PbSO_4 \quad E^0 = +1.682 \text{ V} \tag{20.25}$$

$$PbSO_4/Pb \quad E^0 = -0.359 \text{ V} \tag{20.26}$$

The sum of these potentials yields a value of $E^0 = 2.041$ V. This is not exactly the OCP of a lead–acid battery at open circuit (although it is rather close to it), because the concentration of sulfuric acid is not the standard concentration.

It is clear that the $PbO_2/PbSO_4$ electrode, which serves as the cathode during discharge, is the positive terminal of the battery. In contrast, the $PbSO_4$/Pb electrode, which acts as the anode, is the negative terminal. When the battery is being charged, *the electrodes change roles, but not polarity*. The $PbSO_4$/Pb electrode is now the anode (since $PbSO_4$ is being oxidized to PbO_2), but it is still the positive terminal of the battery. The way in which the potential at each electrode is changed during charge and discharge is shown in Figure 20.7. We must distinguish between two cases. When the battery is in the driving mode, that is – when the battery is the source of energy and it is being *discharged* – the positive terminal is the cathode and the negative terminal is the anode. The same applies also to corrosion and to any other electrochemical process that occurs spontaneously. When the battery is in the driven mode, that is – when it is being *charged* – the positive terminal is the anode and the negative terminal is the cathode. The latter is the case in electroplating, in the electrolytic industry, and in most experiments in the research laboratory, where a current is imposed on the system and the potential is measured, or vice versa.

To summarize Section 20.2, it may be noted that during the first century after the development of the Leclanché cell and the lead–acid battery the field developed steadily but slowly. In the second half of the twentieth century, two major breakthroughs occurred, both related to Li batteries: first, the primary, and then

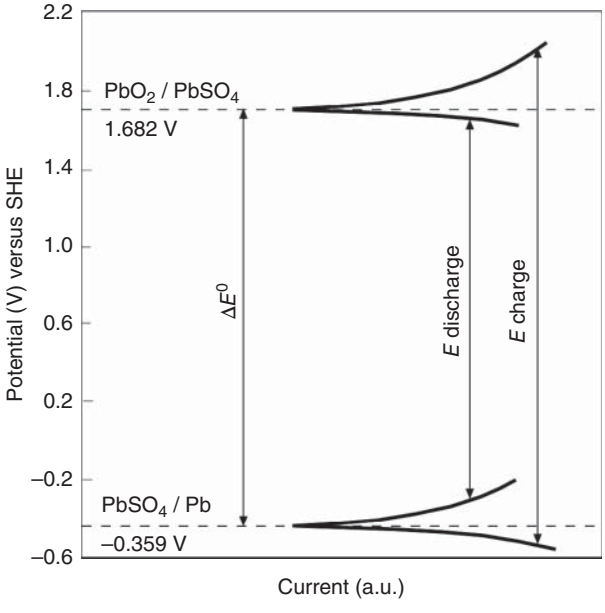


Figure 20.7 The potential of the two electrodes in a lead–acid battery during charge and discharge. Note that the polarity of the cell is not changed. For details see P.J. Moran and E. Gileadi, *J. Chem. Education* **66** (1989) 912.

the rechargeable LIB were introduced. The open-circuit voltage rose from 2.1 V in aqueous systems to more than 4 V in nonaqueous Li batteries, and may reach an OCP of 4.5 V. Considering that the whole electromotive series spans a range of about 6 V (from Li to F_2), this technology may be approaching its limit. It should also be noted that the next challenge is to match the breakthrough in technology with a similar breakthrough in reducing the price, to allow wide use of Li-based batteries.

20.3 Fuel Cells

20.3.1 The Specific Energy of Fuel Cells

As noted in Section 20.2.1, a fuel cell is just a different type of battery, in which the chemical energy is stored in an outside container, rather than inside the battery. This can be a great advantage from the point of view of specific energy since, for extended operation, the weight of the cell itself is insignificant.

The hydrogen–oxygen fuel cell is theoretically the best fuel cell one could devise. The theoretical specific energy is 3.66×10^3 W h kg^{-1}, counting the weight of both hydrogen and oxygen. This should be compared to the value of 208 W h kg^{-1} calculated above for the Ni–Cd. This number is, however, misleading because it is calculated only on the basis of the weight of the fuel and the oxidant, ignoring the weight of the fuel container. The specific energy of molecular hydrogen is the highest of all elements, but in any way of storing it, the weight of the container is at least 10 times the weight of the hydrogen contained in it.

It could be argued that the energy stored in *the fuel cell per se* approaches zero, while the energy of a *fuel-cell system* approaches asymptotically the specific energy of the fuels and their containers. For space applications, both hydrogen and oxygen are carried cryogenically (as the corresponding liquids); hence, the specific energy of the system has to include the containers of both gases. For terrestrial applications, a hydrogen/air fuel cell would be preferred, alleviating the need to store oxygen. Nevertheless, the cost of the electrode materials, the safety aspects of storing hydrogen, and the low density of stored hydrogen, either as a compressed gas or cryogenically, has so far prevented this type of fuel cell from becoming widely used, although it was employed successfully in space missions. There are several types of fuel cells; these will be described in the following sections.

20.3.2 The Phosphoric Acid Fuel Cell (PAFC)

Phosphoric acid fuel cell (PAFC; Figure 20.8a) is a type of fuel cell in which the electrolyte consists of concentrated (85–100%) phosphoric acid. Protons are transported from the anode to the cathode. The two electrodes consist of a mixture of the Pt electrocatalyst supported on carbon black, and a polymeric binder to bind the carbon black particles together, forming an integral structure. A porous carbon paper substrate serves as a structural support for the electrocatalyst layer and as the current collector. The composite structure consisting

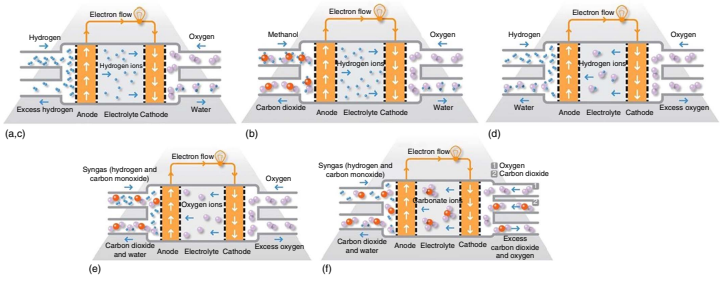


Figure 20.8 The main fuel-cell technology types. (a) Phosphoric acid fuel cell (PAFC). (b) Direct methanol fuel cell (DMFC). (c) Proton exchange membrane fuel cell (PEMFC). (d) Alkaline fuel cell (AFC). (e) Solid oxide fuel cell (SOFC). (f) Molten carbonate fuel cell (MCFC). Source: Reprinted with permission from http://www.fuelcelltoday.com/technologies.

of a carbon black/binder layer on carbon paper substrate forms a three-phase interface, with the electrolyte on one side and the reactant gases on the other side of the carbon paper. A silicon carbide structure holds the phosphoric acid electrolyte. The oxidation and reduction reactions are, respectively,

$$2H_{2(g)} \rightarrow 4H^{+}_{(aq)} + 4e^{-} \tag{20.27}$$

$$O_{2(g)} + 4H^{+}_{(aq)} + 4e^{-} \rightarrow 2H_2O_{(l)} \tag{20.28}$$

Note that the fuel cell operates on H_2 while CO is a poison when present in a concentration greater than 0.5%. If a hydrocarbon such as natural gas is used as a fuel, reforming of the fuel by the reaction

$$CH_{4(g)} + H_2O_{(g)} \rightarrow 3H_{2(g)} + CO_{(g)} \tag{20.29}$$

and shift of the reformate by the reaction

$$CO_{(g)} + H_2O_{(g)} \rightarrow H_{2(g)} + CO_{2(g)} \tag{20.30}$$

are required to generate the required fuel by the cell. Any sulfur compounds present in the fuel have to be removed prior to use in the cell (upstream of the reformer) to a concentration of <0.1 ppmV. The fuel cell itself, however, can tolerate a maximum of 50 ppmV of sulfur compounds. PAFCs are CO_2-tolerant and can even tolerate a CO concentration of about 1.5%, which broadens the choice of fuels they can use. Their electrolyte rejects CO_2; therefore, the gas mixture produced in the steam reformer can be pumped directly into the fuel cell, where H_2 is oxidized at the anode while CO_2 acts as an inert gas, passing through the cell without undergoing any chemical reaction. CO produced as a side product is removed with a so-called *shift reactor*, in which CO is selectively oxidized to CO_2 (see Eq. (20.30)).

Since the amount of catalyst has been reduced to about 0.1 mg cm^{-2} at the anode and 0.2 mg cm^{-2} at the cathode, the cost of the noble metal (Pt) is no longer prohibitive, although it would still constitute a significant factor in the cost of the fuel cell assembly. Moreover, if this type of fuel cell would be widely used (for example, in electric cars using them as the power source), the cost of Pt could increase very much. This cell could be operated at a potential of about 0.70 V, corresponding to an efficiency of 0.70/1.48 = 0.47 (47%),[10] which is comparable to modern conventional power stations. However, the overall system efficiency is only about 40%, since the energy used to operate all the auxiliary equipment, such as pumps and various control units, must be accounted for. The PAFC operates at moderately high temperatures of 150–210 °C. Disadvantages include rather low power density, lower efficiency than other fuel cell types in producing electricity, and aggressive electrolyte.

The PAFC is used in stationary power generators with output in the 100–400 kW range to power commercial premises, and it is also finding

10 For proper comparison with the efficiency of a heat engine, the efficiency of a fuel cell must be calculated on the basis of the heat content of the fuel ($\Delta H/2F = 1.48$ V), not on the basis of the Gibbs energy change ($\Delta G/2F = 1.23$ V).

application in large vehicles such as buses. This fuel cell was the first to be commercialized; until 2001, most fuel cell units sold used PAFC technology.

An interesting possibility proposed was to operate PAFC for the combined purpose of generating electricity and heat. Placing one or more such units in an apartment building could provide electricity, hot water, heating, and cooling locally for each building. Operated in this manner, the efficiency of the system has been estimated to be as high as 80%. Central power stations (both conventional and atomic) also produce large amounts of waste heat, of course, but considering that they are usually built as far from residential areas as possible, using this thermal energy is usually not practical.

20.3.3 The Direct Methanol Fuel Cell (DMFC)

The DMFC (Figure 20.8b) is a relatively recent addition to the family of fuel cell technologies; it was invented and developed in the 1990s by researchers at several institutions in the United States, including NASA and the Jet Propulsion Laboratory. It is similar to the polymer electrolyte membrane fuel cell (PEMFC) in that it uses a polymer membrane as an electrolyte. However, the Pt–Ru catalyst on the DMFC anode is able to draw the hydrogen from liquid methanol, eliminating the need for a fuel reformer. Therefore, pure methanol can be used as fuel; hence the name.

Methanol offers several advantages as a fuel. It is inexpensive but has a relatively high energy density and can be easily transported and stored. It can be supplied to the fuel cell unit from a liquid reservoir that can be kept topped up, or in cartridges that can be quickly changed out when spent.

Methanol can be oxidized electrochemically by a six-electron process, yielding CO_2 and water:

$$CH_3OH + H_2O \rightarrow CO_2 + 6H^+ + 6e^-_M \tag{20.31}$$

The equivalent weight of methanol is only 5.33 g equiv.$^{-1}$. Moreover, methanol is a liquid over the whole temperature range relevant to vehicle operation ($T_m = -97\,^\circ C$, $T_b = +64.7\,^\circ C$). It is easy to manufacture and is inexpensive. The specific energy is lower than that of gasoline (which is itself one third of that of hydrogen), but compared to hydrogen this is compensated for by the fact that the density is high (0.792 g cm^{-3}) and storage presents no problem. Considering that methanol can be manufactured from hydrogen and natural gas, or even from hydrogen and CO_2, it could also be considered as one of the ways of storing hydrogen chemically.

On the other hand, methanol is rather toxic, so why not use ethanol, which is less toxic, has similar physical properties ($T_m = -114.3\,^\circ C$, $T_b = +78.4\,^\circ C$, $\rho = 0.789$ g cm^{-3}), and is also inexpensive? It turns out that ethanol and higher alcohols cannot be oxidized all the way to CO_2, at least not at the potentials relevant to fuel-cell operation. The product of electrochemical oxidation is acetic acid in the case of ethanol and the corresponding higher acid in the case of higher alcohols. It seems that breaking the C—H and the O—H bonds in methanol is easier than breaking the C—C bond in ethanol.

Finally, the choice of DMFC rather than a system operating with a steam reformer that converts methanol to hydrogen and carbon dioxide (as in the PAFC discussed above) is because of the complexity of such systems, which have limited the commercialization of PAFC.

The three main components of any fuel cell are the anode, the cathode, and the membrane. Let us describe them with respect to the DMFC.

(a) *The anode*. Oxidizing methanol is much more difficult than oxidizing hydrogen, and the challenge of finding a good electrocatalyst is greater. Platinum itself is not the best choice, and it seems that an alloy of Pt–Ru is more efficient. The common wisdom at the present time is that adsorption of methanol occurs mostly on Pt sites on the surface of the alloy, and this is also the site of partial oxidation, as shown in the next two equations:

$$\mathrm{Pt} + \mathrm{CH_3OH_{soln}} \rightarrow \mathrm{Pt(CH_3OH)_{ads}} \tag{20.32}$$

$$\mathrm{Pt} - \mathrm{CH_3OH_{ads}} \rightarrow \mathrm{Pt} - \mathrm{CO_{ads}} + 4\mathrm{H^+} + 4\mathrm{e_M^-} \tag{20.33}$$

The oxidation of CO on the Pt site is slow, but it is accelerated by oxygen atoms adsorbed on the Ru site, following the equations

$$\mathrm{Ru} + \mathrm{H_2O} \rightarrow \mathrm{Ru} - \mathrm{OH_{ads}} + \mathrm{H^+} + \mathrm{e_M^-} \tag{20.34}$$

$$\mathrm{Pt\text{-}(CO)_{ads}} + \mathrm{Ru} - \mathrm{OH_{ads}} \rightarrow \mathrm{Pt} + \mathrm{Ru} + \mathrm{CO_2} + \mathrm{H^+} + \mathrm{e_M^-} \tag{20.35}$$

Evidently, this mechanism, if it is indeed the way the system operates, requires that the Pt sites upon which methanol and later CO are adsorbed should be very close, on the atomic scale, to the Ru sites where OH is adsorbed. This requires a very intimate mixing of the two metals, so the methods of preparing the alloy catalyst may be the crucial factor in its effectivity. Alternatively, it may be assumed that there are domains of Pt sites and of Ru sites, and complete oxidation of methanol to CO_2 according to reaction (20.35) only occurs at or very close to the boundaries of such domains. Should this be the case, then the role of nanotechnology in the development of better electrocatalysts (see Chapter 19) becomes evident, and indeed a major effort in the development of better catalysts is conducted in this direction.

There are two avenues of introducing methanol into the cell: either as a fairly concentrated solution in water (about 2–3 M) or as a vapor. Both methods are being studied, each having its own advantages and disadvantages.

(b) *The polymer electrolyte membrane (PEM)*. The solid polymer electrolyte most commonly used in fuel cells is Nafion, which is a derivative of Teflon that has been modified by replacing a certain fraction of the fluorine atoms by sulfonic acid anions $-\mathrm{C} - \mathrm{SO_3^-}$. This serves the purposes of an acid electrolyte, a separator, and a cation-exchange membrane. Sulfonic acid is a strong acid, and its anion rejects CO_2 and other anions, while allowing easy passage of protons across it, which is essential for the proper operation of the cell. The chemical stability of Nafion is very high. Unfortunately, it does not prevent entirely the passage of methanol. If the fuel crosses over and reaches the positive electrode, it is rapidly oxidized, consuming some of the oxygen without

generating any electrical energy. Thus, when methanol is fed into the fuel cell in an aqueous solution, its optimum concentration is determined by the need to minimize mass-transport limitation on the one hand, and charge transfer across the Nafion membrane on the other. This is one of the reasons to use vapor feed of methanol at the anode. It allows dynamic control of the vapor pressure by monitoring the current passed and calculating from it the actual rate of consumption of the fuel, ensuring that most of it will be consumed at the anode, preventing, or at least decreasing, the concentration gradient that would otherwise drive methanol across the membrane.

The worst aspect of the Nafion membrane is its high cost. Massive R&D effort has been going on for years to develop low-cost cation-selective membranes having a performance comparable to that of Nafion, at a much lower price. This is a critical issue because the commercial success or failure of the DMFC hinges on our ability to develop such membranes at a cost of 1–2% of their present price.

(c) *The reduction of molecular oxygen at the cathode.* Oxidation of the fuel is one of the problems of fuel cells. The other is the reduction of oxygen. In an acid solution, the OCP for oxygen reduction should be 1.23 V versus RHE, but in effect it is about 1.0 V. This already reduces the efficiency of energy conversion to about 80%. Moreover, the reduction of oxygen is a relatively slow process, so that further loss of efficiency will result from the activation overpotential. Mass-transport limitation can also be a problem, because of the low solubility of molecular oxygen in water.

In order to be viable, a fuel cell should operate at a high current density, of the order of 1 A cm^{-2}; hence, even a small Ohmic resistance between the two electrodes may cause significant loss of efficiency of energy conversion. This requires optimization of the thickness of the Nafion membrane for low resistivity combined with low permeation of methanol. A further difficulty is that reduction of molecular oxygen could easily lead to the two-electron reduction to hydrogen peroxide instead of the four-electron reduction to water or OH^-. This is to be expected, since the reaction

$$O = O + 2e_M^- + 2H^+ \rightarrow HO - OH \tag{20.36}$$

requires the breaking of a single bond in molecular oxygen, while the reaction

$$O = O + 4e_M^- + 4H^+ \rightarrow 2H_2O \tag{20.37}$$

requires the breaking of two bonds. If the peroxide species is formed as an adsorbed intermediate that is rapidly reduced further to water, this does not pose any problem, but if it is released into the solution, it could decompose to water and oxygen in a disproportionation reaction that proceeds spontaneously in water:

$$2H_2O_2 \rightarrow O_2 + 2H_2O \tag{20.38}$$

In Figure 20.9, the dependence of the rate of oxygen reduction (line 1) and hydrogen oxidation (line 2), as well as the power of a fuel cell (line 3), are shown

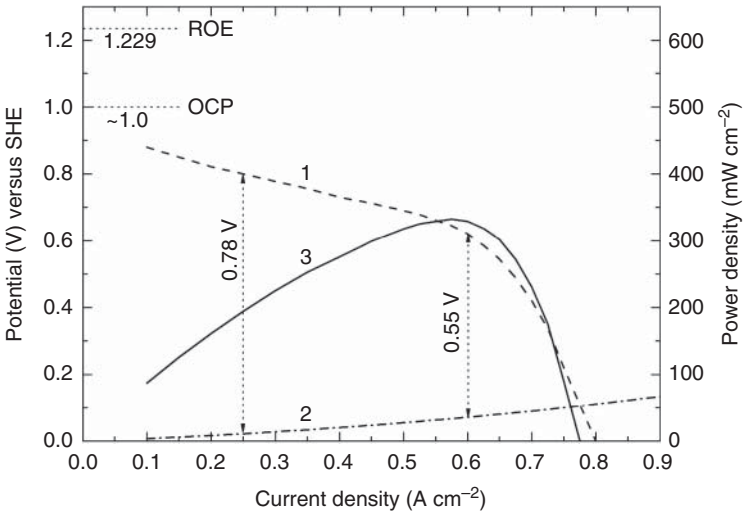


Figure 20.9 Schematic presentation of the current–potential relationship in a fuel cell. Lines 1 and 2 show the polarization curves for oxygen reduction and hydrogen oxidation, respectively. Line 3 shows the power as a function of the current density. Two possible operating points are marked, one for high energy density, the other for maximum power output.

schematically, as a function of the current density. If the application demands the highest energy utilization (as may be the case in a space mission), the cell will be operated at a relatively low current density, as indicated by the vertical line at 0.25 A cm^{-2}. This, however, will require a higher volume and weight, to say nothing of the cost. If the application calls for maximum power output, the device will be operated at a high current density, just before mass-transport limitation sets in, as indicated by the vertical line at 0.60 A cm^{-2} in Figure 20.9.

DMFCs operate in the 60–130 °C temperature range, and are mainly used in applications with modest power requirements, such as mobile electronic devices or chargers and portable power packs. One particular application for DMFCs, which is seeing commercial traction in various countries, is the use of DMFC power units for materials handling vehicles. By switching to fuel cells, warehouses can refuel their trucks in a matter of minutes, compared to the hours it would take to charge a battery. The fuel cells also eliminate the need for a battery charging infrastructure within the warehouse, thereby making more floor space available for other uses.

20.3.4 The Proton Exchange Membrane Fuel Cell (PEMFC)

The proton exchange membrane fuel cell (PEMFC; Figure 20.8c) uses a water-based, acidic polymer membrane as its electrolyte with Pt-based electrodes. This type of fuel cell is also sometimes called *polymer electrolyte membrane fuel cell* (also PEMFC). The anode and cathode are formed on a thin layer on each side of a proton-conducting polymer membrane, used as an electrolyte. In a sense, the electrolyte is composed similarly to the plasticized electrolyte in a LIB, where a liquid electrolyte component is immobilized in a

polymer matrix. It consists of a solid Teflon backbone with a perfluorinated side chain that is terminated with a sulfonic acid group. Hydration of the membrane yields dissociation and solvation of the proton of the acid group. The solvated protons are mobile within the polymer and provide electrolyte conductivity. The membrane has low permeability to oxygen and hydrogen for high coulombic efficiency. A graphite (or metal) plate serves as the plenum for the gas supply and for heat removal. A catalyzed carbon layer is applied to the membrane surface. A thicker gas diffusion layer or porous carbon paper provides gas transport to the reaction zone. The composition and amount of catalyst differ for each electrode, the anode having a lower catalyst loading than the cathode (0.1 versus 0.5 mg cm^{-2}, respectively). The Pt-based catalysts are sensitive to H_2S and CO impurities among others; hence, they must be eliminated from the feedstock for long operation. Water management in the membrane is critical for long-term performance. The proton transport carries water along while carrying the current. The water concentration gradient results in back-diffusion. However, for operation at high current, the anode side of the membrane must be humidified; otherwise, it will dehydrate. By the same token, product water is removed from the cathode side in the air stream to prevent flooding of the active layer.

In PEMFC, hydrogen fuel is processed at the anode where electrons are separated from protons on the surface of a Pt-based catalyst. The protons pass through the membrane to the cathode side of the cell while the electrons travel in an external circuit, generating the electrical output of the cell. On the cathode side, another precious metal electrode combines the protons and electrons with oxygen to produce water, which is expelled as the only waste product; oxygen can be provided in a purified form, or extracted at the electrode directly from the air.

PEMFC cells operate at relatively low temperatures (below 100 °C), are lightweight, and exhibit high power density, high energy conversion efficiency, and compactness; thus, they can tailor electrical output to meet dynamic power requirements. Owing to the relatively low temperatures and the use of precious-metal-based electrodes, these cells must operate on pure hydrogen. The PEMFC was first developed for the Gemini space vehicle. In 1999 Daimler-Chrysler exposed NECAR 4 (NECAR stands for New Electric CAR), a prototype PEMFC-powered car operated with liquid hydrogen. PEMFC cells are currently the leading technology for light-duty vehicles and materials handling vehicles, and to a lesser extent for stationary and other applications.

A variant of the PEMFC that operates at elevated temperatures is known as the high-temperature PEMFC (HT-PEMFC). By changing the electrolyte from being water-based to a mineral-acid-based system, HT-PEMFCs can operate at up to 200 °C. This overcomes some of the current limitations with regard to fuel purity, with HT-PEMFCs being able to process reformate containing small quantities of carbon monoxide. HT-PEMFCs are not superior to low-temperature PEMFCs (LT-PEMFCs); both technologies find niches where their pros are useful. Table 20.1 compares some characteristics of HT-PEMFCs and LT-PEMFCs.

Table 20.1 Comparison between selected specifications of low-temperature and high-temperature PEMFCs.

	LT-PEMFCs	HT-PEMFCs
Operating temperature (°C)	80–100	Max 200
Electrolyte	Water-based	Mineral-acid-based
Pt loading (mg cm^{-2})	0.2–0.8	1.0–2.0
CO tolerance	<50 ppm	1–5 vol.%
Other impurity tolerance	Low	Higher
Power density	Higher	Lower
Cold start?	Yes	No
Water management	Complex	None

Source: Reprinted with permission from http://www.fuelcelltoday.com/technologies.

20.3.5 The Alkaline Fuel Cell (AFC)

The alkaline fuel cell (AFC; Figure 20.8d) is one of the oldest fuel cell types; it was first developed by NASA for the Apollo missions to produce both electricity and water aboard spacecraft. An updated version has been developed and is still in use to provide electrical power for shuttle missions.

AFCs use an alkaline electrolyte such as potassium hydroxide in water and are generally fueled with pure hydrogen. The first AFCs operated between 100 and 250 °C, but typical operating temperatures are now around 70 °C. Because the hydrogen and oxygen kinetics are more facile in alkaline than in acid electrolytes, higher cell voltages can be obtained. Together with the low operating temperature, this permits the use of non-noble metal catalysts, such as Raney Ni, for the fuel electrode. Indeed, Ni is the most commonly used catalyst in AFC units. Silver and spinel-type oxides along with iron phthalocyanines and other porphyrins are good catalysts for the oxygen (air) electrodes. These catalysts cannot be used in acidic electrolytes, as they are soluble in acidic media. Owing to the rate at which the chemical reactions take place, these cells offer relatively high fuel-to-electricity conversion efficiencies, as high as 60% in some applications. The AFC is susceptible to CO_2 contamination of the electrolyte (clogs the pores) when air is used, as well as to poisoning of the Pt and Ni catalysts by sulfide and CO impurities in the feedstock. Therefore, the AFC requires pure hydrogen and oxygen as reactants. Because of the requirement of pure fuels and elimination of CO_2 for long life, terrestrial applications are limited. The AFC technology has been deployed mainly in small stationary systems for backup power.

The AFC cell reactions are as follows:

$$\text{anode:} \quad H_2 + 2OH^- \rightarrow 2H_2O + 2e^- \tag{20.39}$$

$$\text{cathode:} \quad O_2 + H_2O + 2e^- \rightarrow HO_2^- + OH^- \tag{20.40}$$

$$HO_2^- + H_2O + 2e^- \rightarrow 3OH^- \tag{20.41}$$

$$\text{or } O_2 + 2H_2O + 4e^- \rightarrow 4OH^- \tag{20.42}$$

20.3.6 High-Temperature Fuel Cells

Increasing the operating temperature of fuel cells is beneficial in several respects:

(i) The rate of the reaction at both electrodes increases with increasing temperature, decreasing the activation overpotential or effectively eliminating it.
(ii) The conductivity of the ionic conductor (which is in most cases a solid) is increased, thus allowing the passage of high current densities with acceptable resistance overpotential.
(iii) The rate of mass transport is also increased.
(iv) The sensitivity to impurities is reduced or eliminated. For example, PAFC operating at 200 °C could tolerate up to about 2% CO in the gas stream, without significant effects of poisoning. In comparison, a solid oxide fuel cell (SOFC), to be described in this section, could actually use CO as the fuel while operating at 800–1000 °C.
(v) The heat generated in such systems can be used for cogeneration of electricity and heat, thus increasing the overall efficiency of conversion of chemical energy to electricity and heat.

20.3.6.1 The Solid Oxide Fuel Cell (SOFC)

The heart of the high-temperature SOFC (Figure 20.8e) is an yttria-stabilized zirconia (YSZ) film, $Zr_{0.92}Y_{0.08}O_2$, which acts as a solid electrolyte, allowing high conductivity for O^{2-} ions at about 1000 °C. In fact, SOFCs work at the highest temperatures of all fuel cell types (800–1000 °C). The fuel electrode is a porous Ni/ZrO_2 cermet (a term standing for "ceramic-metal"), which serves both as the electrocatalyst and as the current collector. Actually, at such elevated temperatures the reaction kinetics are enhanced, eliminating the need for a metal catalyst, and different fuels, including H_2, regular fossil fuels, and even CO, can be used. The air (oxygen) electrode is a discontinuous layer of strontium-doped lanthanum manganite catalyst, $La(Sr)MnO_3$, coated with a porous doped indium oxide current collector. Cells are connected in series in a *bipolar configuration*, in which the anode of one cell is internally connected to the cathode of the next cell, and only the terminal anode and cathode are connected to the external terminals of the stack. The interconnecting material, which must be a good electronic conductor but at the same time impervious to ions, is manganese-doped lanthanum chromite, $La(Mn)CrO_3$.

The electrolyte is solid, but unlike the PEM, it is dry. This property eliminates many engineering problems of water management, which tend to complicate the design and operation of other types of fuel cells, because the critical temperature of water is 374.15 °C, so it cannot exist in the liquid phase above this temperature. Waste heat is produced at high temperatures and can be used for different purposes, making the overall efficiency of electricity and heat production very high. In fact, the waste heat produced in this type of fuel cells can even be used to produce electricity in a conventional heat engine. SOFCs can have efficiencies of over 60% when converting fuel to electricity; if the heat they produce is also harnessed, their overall efficiency in converting fuel to energy can be over 80%.

The difficulties in the development of SOFCs are in the area of stability of materials rather than in catalysis. Different materials, some of them ionic conductors

with no electronic conductivity and others electronic conductors with no ionic conductivity, must be compatible with each other both chemically at a high temperature and mechanically during temperature cycling. Another constraint is that SOFCs must be shielded to prevent heat loss.

Improvements in materials are steadily made, but the more sophisticated materials developed for this purpose tend to increase the cost. Once the materials issues are overcome, the inherent simplicity of the design and operation may make SOFCs a useful type of electrochemical energy conversion device. Even so, it is an unlikely application for transportation, where the power has to be turned on and off a number of times a day. Cycling the temperature of several tightly connected but different materials, having different heat conductivities and coefficients of thermal expansion, might pose an insurmountable challenge. Moreover, having to start a vehicle by first heating a part to 1000 °C may be time consuming and a waste of energy, particularly considering that cogeneration is not an option in vehicles.

SOFCs are used extensively in large and small stationary power generation. Planar geometries find application in, for example, Bloom Energy's 200 kW off-grid power generators, while SOFCs with output of a few kilowatts are being tested for smaller cogeneration applications, such as domestic combined heat and power (CHP). Bloom's Energy Server® (www.bloomenergy.com), for example, is a distributed power generator that supplies enough power to meet the baseload needs of 160 average homes or an office building. Each Bloom Energy fuel cell is capable of producing about 25 W, enough to power a light bulb. For more power, the cells are sandwiched, along with metal interconnect plates, into a fuel cell "stack." Multiple stacks are aggregated together into a "power module," and then multiple power modules, along with a common fuel input and electrical output are assembled as a complete system. Microtubular SOFCs with output in the watt range are also being developed for small portable chargers. It was estimated that during 2017, 76 MW of SOFC was shipped globally.

20.3.6.2 The Molten Carbonate Fuel Cell (MCFC)

The molten carbonate fuel cell (MCFC; Figure 20.8f) uses a molten carbonate salt suspended in a porous ceramic matrix as the electrolyte. Salts commonly used include lithium carbonate, potassium carbonate, and sodium carbonate. A mixture of alkali carbonates is often chosen to reduce the melting point. The carbonate mixture rejects CO_2 and water, and the temperature is high enough to oxidize CO and most common fuels; hence, this cell is less prone to CO poisoning than lower temperature systems. As a result, MCFC systems can operate on a variety of different fuels, including coal-derived fuel gas, methane, or natural gas, eliminating the need for external reformers. The conductivity of the molten salt is also inherently high. MCFCs operate at high temperature, around 650 °C, and there are several advantages associated with this, for example, it dramatically improves reaction kinetics. Consequently, it is not necessary to boost these cells with a noble metal catalyst, just as in the case of SOFCs.

Disadvantages associated with MCFC units arise from using a liquid electrolyte rather than a solid one, which might pose serious safety hazards, and the requirement to inject CO_2 at the cathode, as carbonate ions are consumed in reactions

occurring at the anode. There have also been some issues with high-temperature corrosion and the corrosive nature of the electrolyte, but these can now be controlled to achieve a practical lifetime.

MCFCs are used in large stationary power generation. Most fuel cell power plants of megawatt capacity use MCFCs, as do large CHP and combined cooling and power (CCP) plants. These fuel cells can work at up to 60% efficiency for fuel-to-electricity conversion, and overall efficiencies can be over 80% in CHP or CCP applications where the process heat is also utilized. Estimated shipments of MCFC amount to 25 MW in 2017.

To conclude Sections 20.3.2–20.3.6, Table 20.2 summarizes selected characteristics of the six different types of fuel cells.

20.3.7 Porous Gas Diffusion Electrodes

There are two ways of increasing the catalytic activity of an electrode in a fuel cell: (i) by finding a better catalyst, which yields a higher value of the exchange current density, and hence a lower overpotential at any chosen current density, and (ii) by increasing the surface area of the electrode per unit *apparent* (geometrical) surface area. Achievements in the former direction have been steady but not spectacular so far. Platinum and some of its alloys are still the best electrocatalysts for oxygen reduction and hydrogen oxidation in acid solutions. The best electrocatalyst at the anode of DMFC seems to be an alloy of Pt–Ru, and most of the research effort has been directed at finding the best composition and methods of preparation of such catalysts.

Development of porous electrodes that have large surface area has come a long way. The effective exchange current density has been increased very significantly, mostly by developing very high-surface-area carbon and involving methods derived from nanotechnology, while the amount of noble metal needed has declined steadily, from about 10 mg cm^{-2} in the early designs to 0.1 mg cm^{-2} for the anode and 0.2 mg cm^{-2} for the cathode in state-of-the-art technology.

However, catalysis is not the only reason for using porous electrodes. For a fuel cell to be economically viable, the current density should be at least 0.5 A cm^{-2}, and preferably higher. Such current densities cannot be reached at planar electrodes even in well-stirred solutions and at substantial concentrations of the electroactive material. In H_2/O_2 fuel cells, both reactants are gases that have low solubility, and vigorous stirring or pumping of the solution is not enough. Moreover, too much of the energy produced by the fuel cell would have to be consumed for this purpose. Porous electrodes are therefore used also to enhance the rate of mass transport.

The way a porous electrode works can best be understood using a single-pore model, shown in Figure 20.10. If one assumes that the pores are conical and are made of a material that is wetted by the solution to some extent, a meniscus will be formed, as shown in Figure 20.10a. An enlarged view of the region of contact between the solution and the gas phase inside a pore is shown in Figure 20.10b. Two factors must be taken into account: the rate of diffusion of the gas through the thin layer of the liquid to the side of the pore that contains the electrocatalyst, and the resistance of the thin layer of the solution, which is determined by the

Table 20.2 Comparison between selected characteristics of the six types of fuel cells.

	AFC	PAFC	DMFC	PEMFC	SOFC	MCFC
System power (kW)	10–100	100–400	10^{-6}–10	10^{-3}–500	$<10^{5}$	300–3000
Cell voltage (V)	1.0	1.1	0.2–0.4	1.1	0.8–1.0	0.7–1.0
Operation temperature (°C)	90–100	150–210	60–130	50–100 (LT); 120–200 (HT)	800–1000	600–700
Catalyst	Ni and others	Pt	Pt–Ru	Pt	None	None
Fuel	Pure H_2	Pure H_2	CH_3OH	Pure H_2	H_2, CH_4, CO, other	H_2, CH_4, CO, other
Efficiency cell (%)	60	>40; co-gen: >85	30–40	50–70	60; co-gen: >80	60; co-gen: >80
Cogeneration	No	Yes	No	No	Yes	Yes
Reformer required?	Yes	Yes	No	Yes	No	No
Electrolyte	Aqueous alkaline, e.g. KOH	85–100% H_3PO_4	Polymer membrane (Nafion)	Polymer membrane. LT: water-based (Nafion). HT: mineral-acid-based (polybenzimidazole, PBI)	Dry YSZ O^{2-} conducting oxide ceramic	Molten alkaline carbonate salt
Number of units shipped in 2017	100	200	2 800	45 500	24 000	0

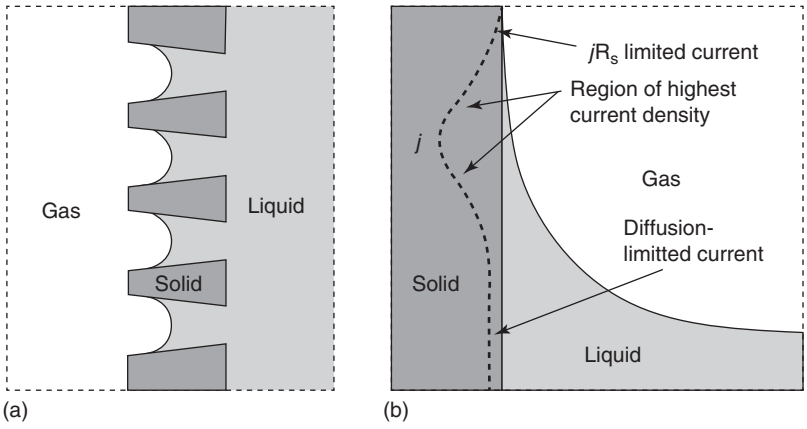


Figure 20.10 (a) The meniscus formed in conical pores of a gas-diffusion electrode. (b) The current density distribution near the three-phase boundary in one of the pores.

cross-section area of the liquid phase that is in intimate contact with the pore. Near the three-phase boundary, the liquid film is very thin, allowing fast diffusion of the gas, but the solution resistance is high. Reaction at the surface of the pore in this region is therefore limited by a high-resistance overpotential. Far away from the edge of the meniscus, the resistance overpotential is low, but here the rate of diffusion is also low, because the layer of liquid through which the reacting gas must diffuse to reach the electrode surface is relatively thick. The reaction will be diffusion-limited in this region. Clearly, the current density must increase first and then decrease as a function of the distance from the edge of the pore, reaching a maximum somewhere in between. This is the most active area of the pore. A great saving in expensive catalyst material can be realized if this area is identified and controlled, and the catalyst material is applied only to it.

It should be emphasized that the single-pore model bears only a distant resemblance to the real structure of a porous electrode, but the principles discussed above are valid. Porous electrodes are usually made by mixing small graphite particles, coated with the catalyst, with a suitable binder. The latter also determines the degree of wettability of the matrix and serves to bind the electrode material to a suitable current collector. This design is necessary to stabilize the gas/liquid interphase, preventing flooding of the electrode on the one hand and drying out on the other. Often a double-pore structure is used, with the larger particles (hence, the larger pores) on the gas phase side of the electrode. This serves to achieve and maintain better stability of the three-phase regions inside the porous structure.

It is important to understand that increasing the roughness factor of a planar electrode increases the rate of *charge transfer*, but has little effect on the rate of *mass transport*. On the other hand, the use of correctly designed porous electrodes can increase the rates of both processes. Thus, the use of porous electrodes will be essential whenever gaseous reactants (e.g. H_2 or O_2) are employed, even after a suitable electrocatalyst will be found.

20.3.8 Fuel-Cell-Driven Vehicles

Undoubtedly, mobile applications have been the major motivation for the huge investments in R&D of fuel cells. Fuel-cell vehicles contribute to the diversification of automobile fuels, emit no CO_2 or environmentally harmful substances during operation, and offer the convenience of gasoline-powered cars. While batteries are improving all the time, and will certainly play an enormous role going forward, limitations in supply, capability, disposal, and charging infrastructure for mobile applications will keep opportunities open for other technologies. A case can also be made that better batteries will improve the viability of fuel-cell hybrid systems, and vice versa.

Mobile applications of interest for fuel cells include forklifts, automobiles, buses, boats, motorcycles, trains, and submarines. The electrical power systems in the Apollo capsules and lunar modules used AFCs. In 1991, the first hydrogen fuel cell automobile was developed by Roger Billings. As of 2015, two fuel cell vehicles were introduced for commercial lease and sale in limited quantities: the Toyota Mirai and the Hyundai ix35 FCEV. The production of Mirai started in 2015; as of December 2017, global sales totaled 5300 units, out of which 2900 units were in the United States and 2100 in Japan. Mirai has a total range of 502 km on a full tank, with a combined city/highway fuel economy rating of 66 mpg-e (3.6 L/100 km).[11] In 2017, sales of the Honda Clarity Fuel Cell began. This car already has a total range of 589 km on a full tank, with a combined city/highway fuel economy rating of 67 mpg-e. Later in 2017, Daimler presented reproduction models of the Mercedes-Benz GLC F-Cell, the world's first EV with fuel-cell/battery powertrain; its LIB can be externally charged by means of plug-in technology. In 2017 Toyota delivered its first two fuel cell electric buses (FCEBs) to Tokyo Metropolitan Government. These buses use a hybrid configuration based on Toyota's Mirai fuel cell system: two 114 kW fuel cell units and a NiMH battery.

Conventional diesel-powered patrol submarines are limited by their battery systems; when these are exhausted, often after a few days, the submarine must surface or use a snorkel to run the main power units to recharge the batteries, placing the vessel at risk. To overcome this limitation, some submarines use PEMFC systems to extend underwater endurance. The German Type 212a, for example, uses nine PEMFCs, providing between 30 and 50 kW each. The ship is silent, giving it an advantage in the detection of other submarines.

In a world of increasingly strict air quality constraints, where major cities such as London are banning diesel buses, heavy duty vehicles have the advantage of better range and faster refueling than battery ones. In some places, it is also easier to build a hydrogen filling station than a similarly sized battery charger, while in others the pure electric solution makes more sense. Toyota, for example, sold its first hydrogen-fuel-cell semi-truck, the Kenworth, in October 2017.

---

11 mpg-e stands for miles per gallon gasoline equivalent. One kilogram of hydrogen is considered as roughly equivalent to one US gallon of gasoline.

20.3.9 Criticism of the Fuel Cells Technology

While the hydrogen/oxygen or hydrogen/air systems are, in principle, the best fuel cells there are, great difficulties of transporting and storing hydrogen have inhibited their use. Efforts have been made to build containers that could store hydrogen at a pressure of 5000 or even 10,000 psi (350–700 atm), replacing the gas tank on conventional vehicles. But, charging a tank with hydrogen at that pressure in a few minutes would generate a large amount of waste heat and would consume one quarter of the energy stored in the hydrogen, or even more. Moreover, the safety hazard of driving a vehicle with hydrogen pressurized at 10,000 psi cannot be ignored.[12]

The low availability of fueling stations is another barrier. As of 2017, there were only 36 hydrogen fueling stations for automobiles publicly available in the United States, the majority of them in California. By May 2017, there were 91 hydrogen fueling stations in Japan. All over China, there are only seven refueling stations. This low availability of fueling stations limits the movement of fuel-cell-driven vehicles.

Some experts believe that hydrogen-fuel-cell cars will never become economically competitive with other technologies, or that it will take decades for them to become profitable. Moreover, fuel cells for use in cars might never be commercially viable because of the inefficiency of producing, transporting, and storing hydrogen and the flammability of the gas, among other reasons. Electricity obtained from hydrogen fuel cells appears to be much more expensive than electricity drawn from the electrical transmission grid, among others due to high energy losses. Furthermore, the large amount of energy required to isolate hydrogen from natural compounds such as water, natural gas, and biomass, package the light gas by compression or liquefaction, transfer the energy carrier to the user, plus the energy lost when it is converted to useful electricity with fuel cells, leaves around 25% for practical use. It has also been argued that fuel-cell vehicles are not really green because of the escape of methane during natural gas extraction and when hydrogen is produced using the steam reforming process.[13]

Fuel-cell technologies such as PEMFC, PAFCs, and MCFCs have all required expensive precious metals, corrosive acids, or hard-to-contain molten materials. Combined with performance that has been only marginally better than alternatives, they have not been able to deliver a strong enough economic value proposition to overcome the status quo. Some makers of legacy fuel-cell energy technologies have tried to overcome these limitations by offering CHP schemes to take advantage of their wasted heat. While CHP does improve the economic value proposition, it only really does so in environments with exactly the right ratios of heat and power requirements on a daily basis. Everywhere else the cost, complexity, and customization of CHP tends to outweigh the benefits.[13]

While fuel cells were invented over a century ago, until now they have not gained widespread adoption because of their inherently high costs and other

12 Note that cars running on natural gas exist, but are not allowed to enter tunnels because of the safety consideration.

13 The Fuel Cell Industry Review 2017, E4tech, November 2017, www.FuelCellIndustryReview.com.

drawbacks. In spite of the huge investments in this field, the return has been very low so far. This can be realized, for example, from the low numbers of units that were shipped during 2017 (see Table 20.2). Several strong players in this arena have already gone into bankruptcy. The United States is letting its power generation tax credit expire, at least temporarily, and has proposed deep cuts to fuel cell research. California has shifted its power generation subsidy largely away from fuel cells toward power storage. If we compare the fuel cells technology to other arenas, such as nanotechnology that arrived later and already has a more significant impact globally, the question arises whether it was justified to put so much efforts and money in the fuel-cell arena.

20.4 Supercapacitors

20.4.1 Electrostatic Considerations

In Chapter 8 (and Chapter 15) we discussed the double-layer capacitance that is always present at the metal/solution interphase. The value of a capacitance, per unit surface area, is given by

$$C = \frac{\varepsilon_0 \varepsilon}{d} \tag{8.3}$$

where $\varepsilon_0 = 8.85419 \times 10^{-12}$ F m^{-1} is the permittivity of free space, and ε is the relative permittivity of the medium between the plates of the capacitor. The important feature of this capacitance for electrochemistry is that it cannot be charged instantaneously. There is a characteristic time constant given by

$$\tau_C = R_S \times C_{dl} \tag{20.43}$$

Thus, when a current step is applied, the potential will approach its final value following the equation

$$E = E_\infty [1 - e^{-t/\tau}] \tag{20.44}$$

It should be noted that the double-layer capacitance is not a capacitor in the usual sense, because it has only one metal plate. The other side of this capacitor is somewhere in solution, depending on the theory assumed and on the composition of the solution. Nevertheless, it behaves as a regular capacitor, except for the fact that its capacitance is often a function of potential.

Capacitors have been used in electronic devices for a century or longer. Their importance hinges on the simple equation

$$Z_C = -\frac{i}{\omega C} \tag{15.9}$$

where $i \equiv \sqrt{-1}$ is the imaginary number and $\omega = 2\pi f$ is the angular velocity. The impedance Z_C approaches infinity for $\omega = 0$, and decreases with increasing frequency. Thus, a capacitor can filter a DC signal, allowing an AC signal to pass, or divert high-frequency noise to the electrical ground, allowing only frequencies below a desired level to be measured.

20.4.2 The Energy Stored in a Capacitor

In this section we shall discuss another property of a capacitor – its ability to store energy. The storage or release of a certain amount of charge in a capacitor, ΔQ, over a range of voltages, $\Delta U = U_{max} - U_{min}$, is expressed using the capacitance as follows:

$$\Delta Q = C \cdot \Delta U \tag{20.45}$$

Thus, if the capacitance is, for example, 10 μF cm^{-2}, the charge needed to create a voltage difference of 1.0 V is 10 μC. The energy consumed in this process, which is the energy stored in the capacitor, is given by

$$W = \frac{CU_{max}^2}{2} = \frac{\Delta Q \cdot U_{max}}{2} \tag{20.46}$$

which in the present example amounts to a very small energy of 5 μJ cm^{-2}. Note, however, that only part of the stored energy is available to applications, because the voltage drop and the time constant over the internal resistance mean that some of the stored charge is inaccessible. The energy density of a capacitor can be increased either by increasing the voltage or by increasing the surface area per unit weight. The former way looks more attractive, because the energy stored is a function of the square of the voltage, but in electrochemical systems it is rather limited. In an aqueous solution one could reach about 0.8–1.2 V. Employing a nonaqueous solvent, perhaps one of those used for Li/thionyl chloride or for Li-ion batteries, could allow a voltage of 3.0–4.0 V, but the double-layer capacitance is usually smaller than in aqueous solutions, so that this approach could probably lead to a factor of about 10 increase in specific energy. On the other hand, increasing the *real* surface area of the electrode has been very successful, and there are porous carbon materials already available, yielding about 2.5×10^3 m^2 g^{-1}. Assuming a capacitance of 10 μF cm^{-2} and a voltage of about 1.0 V leads to a specific energy of 125 kJ kg^{-1} = 34.7 W h kg^{-1}. This is, however, only a theoretical limit (see Section 20.2.2), so it should be compared to a value of about 480 W h kg^{-1} for the LIB based on a $FePO_4$ cathode.

Although the specific energy of supercapacitors is insufficient compared with batteries, capacitors have the important advantage of the specific power. Specific power describes the speed at which energy can be delivered to/absorbed from the load. The maximum power is given by

$$P_{max} = \frac{U^2}{4R_i} \tag{20.47}$$

where U is the applied voltage and R_i is the internal DC resistance of the capacitor. The described maximum power P_{max} specifies the power of a theoretical rectangular single maximum current peak of a given voltage. In real circuits, the current peak is not rectangular and the voltage is smaller, caused by the voltage drop. Thus, the IEC 62391–2 standard has established a more realistic effective power P_{eff} for supercapacitors for power applications, which is half P_{max}.

20.4.3 The Essence of Supercapacitors

Supercapacitors, also known as *electrochemical capacitors* and *ultracapacitors*, are a family of electrochemical energy-storage devices that are ideally suited to the rapid storage and release of energy. The term "supercapacitor" was coined in 1999 by Brian Evans Conway. Unlike ordinary capacitors, supercapacitors do not use a solid dielectric to separate the charges. Instead, they use electrostatic double-layer capacitance and electrochemical pseudocapacitance,[14] both of which contribute to the total capacitance of the capacitor. Compared with conventional capacitors, the specific energy of supercapacitors is several orders of magnitude higher (hence, the "super" prefix). Supercapacitors also have a higher specific power than most batteries, but their specific energy is somewhat lower. Through appropriate cell design, both the specific energy and specific power ranges for supercapacitors can cover several orders of magnitude, and this makes them extremely versatile as a stand-alone energy supply, or in combination with batteries as a hybrid system.

Supercapacitors consist of two metal foils (current collectors), each coated with an electrode material such as activated carbon. The foils serve as the power connection between the electrode material and the external terminals of the capacitor. The electrode is typically made of porous, spongy material with an extraordinarily high specific surface area (SSA), such as activated carbon. The electrodes are kept apart by an ion-permeable membrane (separator)[15] used as an insulator to protect the electrodes against short-circuits. This construction is subsequently rolled or folded into a cylindrical or rectangular shape, and can be stacked in an aluminum can or an adaptable rectangular housing. Then, the cell is impregnated with a liquid or viscous electrolyte of organic or aqueous type. The electrolyte, an ionic conductor, enters the pores of the electrodes and serves as the conductive connection between the electrodes across the separator. Finally, the housing is hermetically sealed to ensure stable behavior over the specified lifetime.

Figure 20.11 shows the charge distribution in a supercapacitor both in the charged and in the discharged states. When the electrodes are polarized, ions in the electrolyte form electric double layers of opposite polarity to the electrode's polarity. For example, positively polarized electrodes will have a layer of anions at the electrode/electrolyte interphase along with a charge-balancing layer of cations adsorbing onto the negative layer. The opposite is true for the negatively polarized electrode. Additionally, depending on the electrode's material and surface shape, some ions may permeate the double layer, become specifically adsorbed ions, and contribute with pseudocapacitance to the total capacitance of the supercapacitor.

---

14 The double-layer capacitance represents storage of the electrical energy achieved by separation of charge in the Helmholtz double layer. Pseudocapacitance represents storage of the electrical energy achieved by faradaic redox reactions with charge transfer.

15 Open capacitor papers, nonwoven porous polymeric films such as polyacrylonitrile or Kapton, woven glass fibers, and porous woven ceramic fibers have been used as separators.

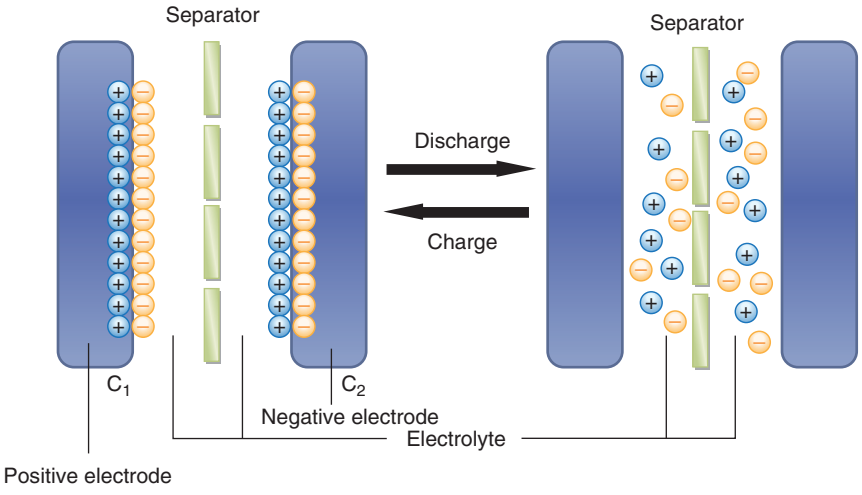


Figure 20.11 The charge distribution in an electrochemical double-layer capacitor (EDLC) in the charged (left) and discharged (right) states. Source: Reproduced with permission from https://www.sciencedirect.com/science/article/pii/S2211285512001826.

The two electrodes form a series circuit of two individual capacitors C_1 and C_2. The total capacitance, C_{total}, is given by

$$C_{\text{total}} = \frac{C_1 \cdot C_2}{C_1 + C_2} \tag{20.48}$$

Supercapacitors may have either symmetric or asymmetric electrodes. Symmetry implies that both electrodes have the same capacitance value, yielding $C_{\text{total}} = 0.5C_1 = 0.5C_2$. For asymmetric capacitors, the total capacitance can be taken as that of the electrode with the smaller capacitance (i.e. if $C_1 \gg C_2$, then $C_{\text{total}} \approx C_2$).

Supercapacitors can store by several orders of magnitude more energy per unit mass or volume than a conventional capacitor because (i) charge separation takes place across a very small distance in the electrical double layer that constitutes the interphase between an electrode and the adjacent electrolyte; (ii) an increased amount of charge can be stored on the electrode thanks to its high surface area, which is established by a large number of pores. The mechanism of energy storage is inherently rapid because it simply involves movement of ions to and from electrode surfaces. In batteries, additional steps, such as heterogeneous charge transfer and chemical phase changes, introduce relatively slow steps into the process of energy storage and delivery. For similar reasons, supercapacitors exhibit a very high degree of reversibility in repetitive charge–discharge cycling.

Supercapacitors are divided into three family members, based on the design of the electrodes and the storage principle:

1. *Electrochemical double-layer capacitors (EDLCs)*, with activated carbon electrodes or derivatives with much higher electrostatic double-layer capacitance than electrochemical faradaic pseudocapacitance. Since EDLCs are

not limited by the electrochemical charge transfer kinetics of batteries, they can charge and discharge at a much higher rate, with lifetimes of more than one million cycles. The EDLC energy density is determined by the operating voltage and the specific capacitance (in units of $F\,g^{-1}$ or $F\,cm^{-3}$) of the electrode/electrolyte system. The specific capacitance is related to the SSA accessible by the electrolyte, its interfacial double-layer capacitance, and the electrode material density. Electrode materials include activated carbon, carbide-derived carbon, carbon aerogel, graphene, and CNTs.

2. *Pseudocapacitors* have a much higher faradaic pseudocapacitance than static double-layer capacitance. They store electrical energy by electrosorption, reduction–oxidation (redox) reactions, and intercalation processes. The pseudocapacitance may be higher by a factor of 100 than the double-layer capacitance for the same electrode surface. Unlike EDLCs, where the electrical charge is stored electrostatically with no interaction between the electrode and the ions, a pseudocapacitor involves a chemical reaction at the electrode. Pseudocapacitance is accompanied by an electron charge transfer between electrolyte and electrode coming from a de-solvated and adsorbed ion. The amount of electric charge stored in a pseudocapacitor is linearly proportional to the applied voltage. Electrochemical pseudocapacitors use transition-metal oxide or conductive polymer electrodes with a high amount of electrochemical pseudocapacitance. Electrode materials include MnO_2, IrO_2, mesoporous films of α-MoO_3, V_2O_5 thin films on CNTs, conducting polymers such as polypyrrole (PPy), polyaniline, or derivatives of polythiophene such as poly(3,4-ethylenedioxythiophene) (PEDOT), etc. Ruthenium dioxide (RuO_2) has been commonly used as it introduces a redox couple between two valence states of ruthenium, to result in higher capacitance for carbon electrodes (but with a penalty in slower time constant for reacting to a pulse demand).
3. *Hybrid capacitors*, such as Li-ion capacitors, use asymmetric electrodes, one of which exhibits mostly electrostatic and the other mostly electrochemical capacitance. The device will be half capacitor and half battery. Therefore, the term "*supercapattery*" has become a generic term for various hybrid devices combining the merits of rechargeable battery and supercapacitor. In this context, it should be noted that each type of battery can be designed in different ways to maximize either the energy density or the power density. Indeed, there are systems that contain two batteries: one with high energy but low power density, charging the other battery that has a high power density but low energy density, which in effect plays the role of a supercapacitor. One of the advantages of hybrid supercapacitors is that they can yield much higher specific capacitance and energy density than symmetric supercapacitors. One example of electrode materials is a nickel hydroxide cathode with a carbon anode.

20.4.4 Advantages of Supercapacitors

Compared to other energy storage systems such as batteries and fuel cells, supercapacitors are superior in terms of life expectancy, reversibility, power

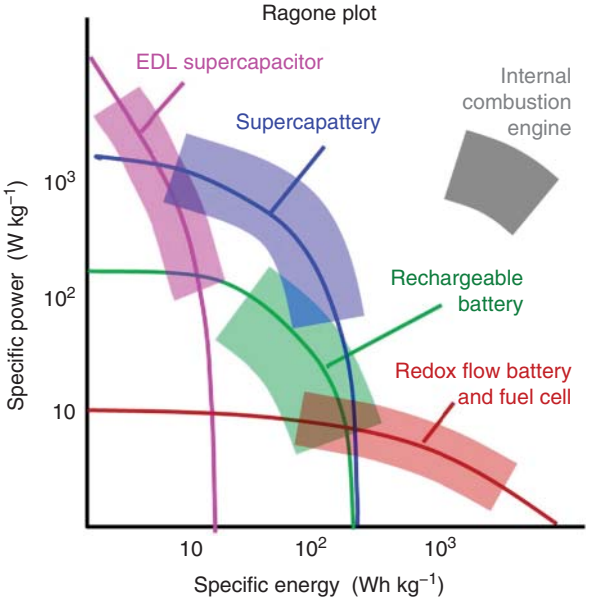


Figure 20.12 A schematic Ragone plot showing the position of EDLCs and supercapatteries relative to other energy technologies. Source: Reproduced with permission from G.Z. Chen, *Intern. Mater. Rev.*, **62** (2017) 173.

density, shelf life, efficiency, operating temperatures, environmental friendliness, and safety. These will be explained here.

(1) *High specific power*. Supercapacitors exhibit a much higher specific power (1–10 kW kg^{-1}) when compared to Li-ion batteries. Figure 20.12 shows the so-called "Ragone plots," comparing the practically achieved specific power and specific energy of EDLCs and supercapattery to those of batteries, fuel cells, and ICEs.

(2) *Instantaneous charge/discharge*. Since a supercapacitor stores electrical charges both at the electrode surface and in the bulk near the surface of the solid electrode, rather than within the entire electrode, the charge–discharge reaction will not necessarily be limited by ionic conduction into the electrode bulk, so the charging and discharging rates are much faster than the electrochemical redox reactions inside batteries. It would be nice to think that charging and discharging an electrochemical supercapacitor is "instantaneous," but this is not quite the case. Recalling Eq. (20.43), we note that the characteristic time is a product of the capacitance and the associated solution resistance. It is hard to calculate a priori the solution resistance of a highly porous electrode, and there may be a range of resistances in a given capacitor. Assuming, for the sake of demonstration, an effective resistance of 0.10 Ω combined with a capacitance of 250 F would lead to a time constant $\tau = 25$ seconds. Inserting this value into Eq. (20.44) leads to a discharge from full charge to 10% in one minute. In the nomenclature used for batteries, this corresponds to a discharge rate of 60 C. No battery is built to survive such high rates of discharge. The characteristic time for discharge could be tailor-made to fit different applications. It may be hard to shorten it for a

very large capacitor below 60 seconds, but it is always possible to slow it down as desired.

(3) *Cycle life*. When a battery is cycled between its charged and discharged states, chemical reactions occur at both electrodes. These reactions occur in the solid state, and in most cases involve a change in the volume of the species in the transition from one state to the other. While the battery may have been designed and even built to have an optimal structure (particle size, crystal structure, bonding, etc.), this may change just a little bit in each cycle, but after a thousand cycles the changes may accumulate to the point that the battery can no longer operate well, or that its charge capacity may decline significantly. Such aging processes are inherent to all types of rechargeable batteries, although some may last more cycles than others. In contrast, when an electrochemical supercapacitor is charged and discharged, this is merely an electrostatic process of creating an excess or deficiency of electrons on one side of the capacitor, and an inverse excess or deficiency of ions at the solution side of the interphase. No chemical reaction takes place and no changes in volume or structure occur. Consequently, while LIB life may extend from 1000 to 10,000 cycles and their life expectancy is estimated at 5–10 years, supercapacitors can survive up to about one million cycles and their life expectancy is estimated at up to 30 years.

(4) *Long shelf life*. Most rechargeable batteries, if left on the shelf unused for months, will degrade and become useless due to self-discharge and corrosion. In contrast, supercapacitors maintain their capacitance and thus are capable of being recharged to their original condition, although self-discharge over a period of time can lead to a lower voltage. It has been reported that electrochemical supercapacitors can sit unused for several years but still remain close to their original condition.

(5) *High efficiency*. Electrochemical supercapacitors are reversible with respect to charging and discharging throughout their complete operating range of voltage, and the energy loss to heat during each cycle is relatively small and readily removed (i.e. heat management is easy). This means that the cycle efficiency of electrochemical supercapacitors is high (90–95%) even when operating at rates above 1 kW kg^{-1}.

(6) *Wide range of operating temperatures*. Electrochemical supercapacitors can function effectively at extremely high and low temperatures. Their typical operating temperature ranges from −40 to 70 °C. This is advantageous for military applications, where reliable energy storage is required to run proprietary electronic devices under any temperature conditions.

(7) *Environmental friendliness*. Electrochemical supercapacitors, in general, do not contain hazardous or toxic materials, and their waste materials are easily disposed.

(8) *Safety*. In normal circumstances, electrochemical supercapacitors are much safer than batteries, in particular Li-ion batteries.

20.4.5 Barriers for Supercapacitors

Although electrochemical supercapacitors have many advantages over batteries and fuel cells, they also face some challenges at the current stage of technology:

(1) *Low specific energy*. The calculation at the end of Section 20.4.2 shows that the theoretical specific energy of an electrochemical supercapacitor is only about 7% of that of a LIB. Commercially available electrochemical supercapacitors can provide specific energy of only 3–4 W h kg^{-1}.
(2) *High cost*. The costs of raw materials and manufacturing continue to be major challenges for electrochemical supercapacitor commercialization. The main cost of a supercapacitor arises from its electrode materials. At present, for practical purposes carbon and RuO_2 are the most common electrode materials used in commercial supercapacitors. However, carbon materials, in particular those with a high surface area, are presently expensive (US\$50–100 kg^{-1}), not to mention the cost of a rare metal oxide such as RuO_2. In addition, the separator and the electrolyte can also boost the expense. For example, if the supercapacitor uses organic electrolytes, their cost is far from being negligible.

20.4.6 Applications of Supercapacitors

The first electrochemical supercapacitor patent was filed in 1957. However, it was not until the 1990s that the supercapacitors technology began to draw some attention, in the field of HEVs. It was found that the main function of a supercapacitor could be to boost the battery or fuel cell in a HEV to provide the necessary power for acceleration, with an additional function being to recuperate brake energy. Further developments have led to the recognition that supercapacitors can play an important role in complementing batteries or fuel cells in their energy storage functions by providing backup power supplies to protect against power disruptions. As a result, the US Department of Energy has designated electrochemical supercapacitors to be as important as batteries for future energy-storage systems. Many other governments and enterprises have also invested time and money into exploring, researching, and developing such technologies. Several standards already exist, including IEC/EN 62391–1, IEC 62391–2, IEC 62576, and BS/EN 61881–3.

Supercapacitors are used in applications requiring many rapid charge/discharge cycles rather than long-term compact energy storage. Applications include EVs, trams, metro trains, elevators, powering emergency actuators for doors and evacuation slides used in airliners, consumer electronics, defibrillators for medical use, grid power buffers, uninterruptible power supplies (UPSs), pulse laser techniques, street lights, storage of the energy generated by solar cells, hybrid batteries, and military applications (e.g. motor startup for large engines in tanks and submarines, and phased array radar antennae). Combining supercapacitors with batteries can yield improved performance in HEVs, including powerful acceleration, braking energy recovery, excellent cold weather starting, and increased battery life.

20.5 Hydrogen Storage

Interest in using hydrogen as a secondary energy carrier was revived by the desire to reduce environmental pollution due to automobile emissions, and in an effort to reduce the use of fossil fuel in general. Hydrogen is in some respects the ideal fuel:

(1) *Availability*. Hydrogen is the most common element in the Universe, roughly 74 wt% or 92 at.%. The amount of free hydrogen on Earth is negligible, however; less than 1% of it is present as pure molecular hydrogen gas (H_2), while the overwhelming majority is chemically bound as protons in water, and some is bound to liquid or gaseous hydrocarbons. Hydrogen is renewable in unlimited amounts (provided there is a method of producing energy that does not depend on burning fossil fuel, e.g. solar, wind, hydroelectric, or nuclear).
(2) *Ease of production*. Hydrogen can easily be separated from oxygen in water via electrolysis. This process currently has an efficiency of about 67% (e.g. the maximum in 2014 for PEM electrolyzers).
(3) *It is clean, leaving only water as the product of its oxidation*. The maximum efficiency of the hydrogen–oxygen reaction is achieved via electrochemical conversion of hydrogen and oxygen to current, heat, and water in fuel cells, via cold combustion, and not via direct burning of hydrogen and oxygen to steam.
(4) *It is the lightest element*. Because of its low equivalent weight, it has the highest specific energy (38 kW h kg^{-1}, compared to 14 kW h kg^{-1} in gasoline).
(5) *Feasible transportation*. For hydrogen transport from the production site over long distances (thousands of kilometers), pipelines are thought to be feasible; they would carry energy much more efficiently than high-voltage, direct-current power lines. Yet, transport via pipelines will not be necessary or economically attractive until a regenerative hydrogen economy has been established on a large scale, entailing very large amounts of energy for transport. A typical pipeline may have a diameter of 1.6–1.8 m and supply 23 GW of H_2. The transport costs for a distance of 3000 km would be around € 0.015 kW h^{-1} of hydrogen, taking into account energy losses of 15% due to the compression and transport of the gas.

Unfortunately, much of these apparent advantages are lost as a result of the difficulty of storage and transportation. In reality, one should consider the *effective equivalent weight* of hydrogen, which is the weight of 1 g stored, including the weight of the medium in which it is stored. In the chemical industry, hydrogen is usually stored either as compressed gas at high pressure, or as cryogenic liquid below its boiling point of −252.87 °C. Both these methods are expensive and consume a lot of energy. For example, the amount of energy needed to compress hydrogen to 700 atm can be as high as 15–25% of the energy in the stored hydrogen. Cryogenic storage consumes at least 30% of the energy stored. In addition, the weight of the container, particularly for compressed hydrogen, could increase its effective equivalent weight significantly. Furthermore, storage requires careful and costly safety equipment.

The low density of hydrogen at ambient temperature (0.089 kg m^{-3}) results in a low energy per unit volume. In order to run a vehicle for 500 km between two fills, the mass of hydrogen required is 5–13 kg, depending on driving cycle and vehicle type. Hence, the space required to store that much hydrogen is unimaginably high as compared to today's tank sizes of gasoline-driven vehicle. The implication of this is that either large compression is required to store hydrogen more effectively, or liquid- or solid-state storage is used. The US DOE has set tough targets for hydrogen storage for the development of hydrogen-fueled vehicles (https://energy.gov/eere/fuelcells/doe-technical-targets-onboard-hydrogen-storage-light-duty-vehicles). The ultimate values include a net useful specific energy of 2.2 kW h kg^{-1} (or 0.065 kg-H_2 kg-system^{-1}); a net useful energy density of 1.7 kW h L^{-1} (or 0.050 kg-H_2 L-system^{-1}); fuel cost of US$8 kW^{-1} h^{-1} net (US$266 kg-H_2^{-1}); operability between −40 and 60 °C (sun); 1500 cycles from quarter tank to full; minimum delivery pressure from storage system – 4.9 atm in fuel cells; maximum delivery pressure from storage system – 11.8 atm; charging/discharging rates (system fill time, per SAE J2601) – 3 to 5 minutes; and fuel purity meets or exceeds SAE J2719. These targets show the importance of both gravimetric and volumetric capacity of the store, as one does not want the store to be too heavy (limiting the range of the vehicle) or too voluminous (limiting the luggage space). Figure 20.13 illustrates the DOE targets for hydrogen storage.

It should be noted that often the costs associated with the production and storage of hydrogen are overlooked. Electrolysis of water requires considerable energy; its cost thus largely depends on energy costs. When using salty water, it has to be distilled by reverse osmosis. As electrolysis and desalination systems require large amounts of energy, a small hydroelectric power (SHP) plant is to

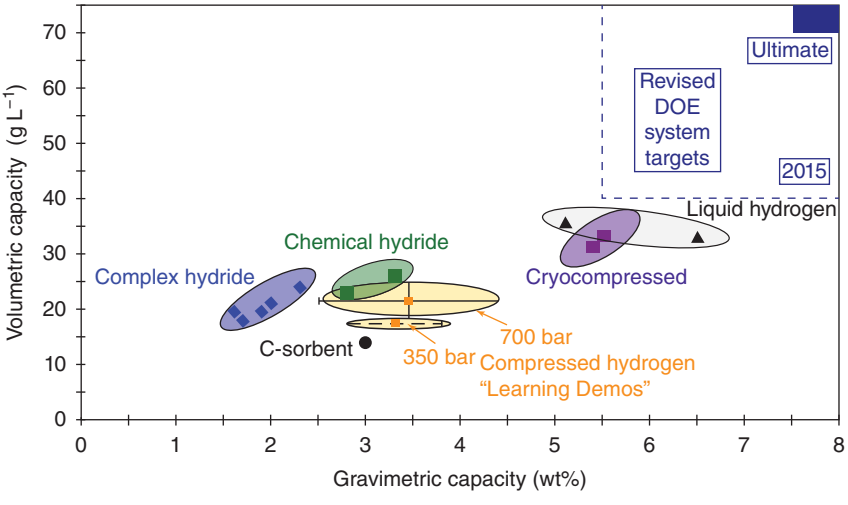


Figure 20.13 Volumetric capacity versus gravimetric capacity of available hydrogen storage technologies, as drawn by the U.S. DOE. Source: Reprinted with permission from https://energy.gov/eere/fuelcells/status-hydrogen-storage-technologies.

be constructed in order to reduce costs. Energy costs for SHP should be taken into account too. In order to obtain liquid hydrogen, gas hydrogen is compressed under high pressure, and then the compressed gas is cooled with liquid nitrogen, and finally is dilated in turbines. As the liquefaction process is realized by cooling the gas to the point where it becomes liquid, compressors, heat exchangers, turbine, and reduction valves are required.

The liquefaction process requires approximately 30% of hydrogen's ignite energy in electrical energy form. The main problem in the production and storage of liquid hydrogen is loss due to vaporization; as the capacity of the storage tank increases, the losses diminish. Vacuum is used for isolation purposes between the inner and outer walls of the tank. Liquid hydrogen can be stored in small and large quantities, but cylindrical tanks for natural gas are not feasible as this leads to large amounts of vaporization losses. Storage of compressed hydrogen gas is the most applied and simplest storage type, as it only requires a compressor and compression tank. On the other hand, high levels of compression mean higher investment and operation costs, and this storage type is not cost-effective for storing large quantities of hydrogen.

Hydrogen burns cleanly in air, producing only water. It has a wide flammability range. Since the lower explosive limit of hydrogen in air (13 vol.%) is higher than the lower flammability limit, hydrogen generally burns rather than explodes. The flame velocity of hydrogen is much higher than that of conventional gaseous fuels. When mixed with pure oxygen in a 2 : 1 atomic ratio of hydrogen to oxygen and ignited, hydrogen detonates violently since there is no inert nitrogen to slow down the wave front. The energy required to ignite a hydrogen air mixture is very low (~0.02 MJ), about 1/14 of the energy needed to ignite natural gas. The combination of all physical properties of hydrogen is unique among all fuels, and this combination has safety implications for the use of hydrogen in bulk. However, the low density and high diffusivity of hydrogen in air means that in case of any leakage, hydrogen will disperse off very fast. Besides, hydrogen has very high net calorific value, with a lower heating value (LHV)[16] of 120 MJ kg^{-1}, much higher than that of petrol (43 MJ kg^{-1}).

Hydrogen can be used for power generation either by combustion in an ICE, producing mechanical power, or electrochemically by using a fuel cell, producing electrical power. In both cases, hydrogen reacts with oxygen to produce water. While in the case of fuel cells this is the only product, the combustion in ICE is at very high temperatures, which might result in the formation of NOX as well.

Hydrogen is commonly found in combination with other elements, in the form of either hydrocarbons or water. Hydrogen can be produced from these by electrolysis, reforming, fermentation, etc. If a renewable source is used, hydrogen production can be free of CO_2 emissions, but if produced from fossil fuels – then it requires CO_2 capture and sequestration. Hydrogen is thus a

16 The lower heating value of a fuel, also known as net calorific value, is defined as the amount of heat released by combusting a specified quantity (initially at 25 °C) and returning the temperature of the combustion products to 150 °C, which assumes that the latent heat of vaporization of water in the reaction products is not recovered.

versatile energy carrier and is seen to play a major role in solution to lowering the CO_2 emissions. Therefore, one sector where hydrogen could have a major impact is transportation.

There is thus a clear incentive to develop better methods to store hydrogen, which have potential for higher energy density and will permit the gas to be used at close to room temperature, and at moderate pressures. Current hydrogen storage approaches involve low-pressure liquid hydrogen tanks (near the boiling point of hydrogen), compressed hydrogen gas tanks (where hydrogen is typically stored at near-ambient temperatures), cold or cryogenic compressed hydrogen (sub-ambient temperatures greater than 150 K or 150 K and below, respectively), metal hydrides, high-surface-area adsorbents, and chemical hydrogen storage materials. All of these, except chemical hydrogen storage, allow "reversible" on-board hydrogen storage because hydrogen regeneration or refill can take place on-board the vehicle. In contrast, in the case of chemical hydrogen storage approaches, such as a chemical reaction on-board the vehicle to produce hydrogen, hydrogen regeneration is not possible on-board the vehicle, and thus, these spent materials must be removed from the vehicle and regenerated off-board.

Hydrogen can be stored physically as either a gas or a liquid. Storage of hydrogen as a gas typically requires high-pressure tanks (350–700 atm). Storage of hydrogen as a liquid requires cryogenic temperatures because the boiling point of hydrogen at 1 atm is −252.8 °C. Liquid hydrogen has a low density of 70.8 kg m^{-3} (7% of that of water).

Many metals and alloys are capable of reversibly absorbing large amounts of hydrogen. Charging can be done using molecular hydrogen gas or hydrogen atoms from an electrolyte. Molecular hydrogen is dissociated at the surface before absorption; two H atoms recombine to H_2 in the desorption process. The thermodynamic aspects of hydride formation from gaseous hydrogen are described by the pressure–composition isotherms (Figure 20.14). The host metal initially dissolves some hydrogen as a solid solution (α-phase). As the hydrogen pressure and the concentration of H in the metal are increased, local interactions between hydrogen atoms become important, and nucleation and growth of the hydride (β) phase begins. While the two phases coexist, the isotherms show a flat plateau (according to the Gibbs' phase rule), the length of which determines how much H_2 can be stored reversibly with small pressure variations. In the pure β-phase, the H_2 pressure rises steeply with the concentration. At higher H_2 pressure, more plateau and more hydride phases may be formed. The two-phase region ends in a critical point T_C, above which the transition from α- to β-phase is continuous. The plateau or equilibrium pressure $p_{eq}(T)$ depends strongly on temperature, and is related to the changes ΔH and ΔS of enthalpy and entropy, respectively, by the van't Hoff equation:

$$\ln\left(\frac{p_{eq}}{p^0_{eq}}\right) = -\frac{\Delta H}{R}\frac{1}{T} + \frac{\Delta S}{R} \tag{20.49}$$

As the entropy change corresponds mostly to the change from molecular hydrogen gas to dissolved hydrogen, it is roughly 130 J K^{-1} mol^{-1} for all metal–hydrogen systems under consideration. The enthalpy term characterizes

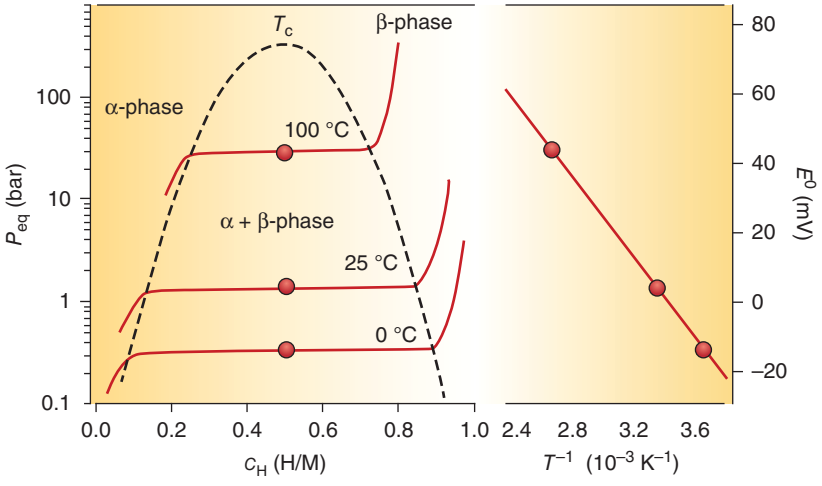


Figure 20.14 Pressure–concentration–temperature plot and a van't Hoff curve (logarithm of the equilibrium or plateau pressure against the reciprocal temperature). Values are for $LaNi_5$. The vertical axes indicate the corresponding hydrogen pressure or the equivalent electrochemical potential. From the slope of the van't Hoff plot, experimental values of the enthalpy of hydride formation ΔH can be extracted. Source: Reproduced with permission from L. Schlapbach and A. Züttel, *Nature* **414** (2001) 353.

the stability of the metal–hydrogen bond. To reach an equilibrium pressure of 1 atm at 27 °C, ΔH should amount to 19.6 kJ mol-H^{-1}. The operating temperature of a metal hydride system is fixed by the plateau pressure in thermodynamic equilibrium and by the overall reaction kinetics (L. Schlapbach and A. Züttel, 2001).

The reversible storage of hydrogen in the form of *metal hydride* has several advantages over conventional gaseous and liquid H_2 storage. Hydrides offer pronounced volumetric density compared to compressed gas. In fact, hydrogen densities achievable with hydride containers approach those of liquid storage, a result of the fact that the hydrogen packing density within a given hydride crystal is usually significantly higher than that of liquid hydrogen. Hydrides also give excellent insurance that the hydrogen released is of very high purity. Reactive species, such as O_2, H_2O, and CO, tend to be gettered by the hydride, and inert species, such as N_2 and CH_4, can be purged away during the first few percentages of discharge. Thus, hydride storage units are often used in special applications where extremely high-purity H_2 is required (e.g. as gas chromatograph carrier gas or for high-purity furnace atmospheres). In addition, because of the low pressures and near-ambient temperature involved in the use of metal hydrides, they are generally viewed to be a safer means of handling and storing hydrogen than cryogenic liquid and compressed gas. The main drawbacks of hydride storage, however, are relatively low hydrogen weight percentages, slow reaction kinetics, low reversibility, high dehydrogenation temperatures, high cost of the intermetallic compounds, and the need to invest in heat exchangers for the storage containers.

Which metallic systems are appropriate for hydrogen storage? We can differentiate between interstitial hydrides (e.g. $LaNi_5H_6$) and complex hydrides

(e.g. $NaAlH_4$). *Interstitial hydrides* most commonly exist within metals or alloys, and their bonding is generally considered metallic. These systems are usually nonstoichiometric, with variable amounts of hydrogen atoms in the lattice. They readily absorb and discharge hydrogen at room temperature and atmospheric pressure. The reaction is simple, without any intermediate compounds, but with significant volume change (~20%). Unfortunately, their application is still limited, as they are capable of storing only 1–2 wt% hydrogen, insufficient for automotive applications. For example, hydrogen could be stored in palladium as $PdH_{0.6}$ at low pressure, and be easily retrieved from it, but this is not practical because its effective equivalent weight in this medium is 177 (0.56 wt% hydrogen), and the price of palladium is prohibitive. When the hydride consists of two metals, the term *intermetallic hydride* is often used, and its formula is $A_mB_nH_x$, where A is an element with a high affinity to hydrogen and B is an element with a low affinity to hydrogen. B is often at least partially Ni, since it is an excellent catalyst for hydrogen dissociation. The properties of the intermetallic compounds are determined by the interaction between the interstitial hydrogen atoms and the metal atoms and therefore, they are largely dependent on the crystal structure of the compounds. AB_5 (e.g. $LaNi_5$), AB_2 (e.g. ZrV_2, $ZrMn_2$ and $TiMn_2$), AB_3 (e.g. $CeNi_3$ and YFe_3), A_2B_7 (e.g. Y_2Ni_7 and Th_2Fe_7), A_6B_{23} (e.g. Y_6Fe_{23}), AB (e.g. TiFe and ZrNi), A_2B (e.g. Mg_2Ni and Ti_2Ni), and vanadium-based solid-solution alloys are among the important types of intermetallic compounds. It has been shown that AB_5, AB_2, and A_2B-type alloys have excellent hydrogen-absorbing properties. For example, an alloy of iron with titanium can absorb and release hydrogen reversibly. With a composition of $TiFeH_{1.95}$ at maximum loading, it stores 1.89 wt% hydrogen, not including the weight of the container in which it is stored. Volumetrically, it stores 0.1008 kg L^{-1} hydrogen. Another alloy developed for the same purpose is made of alloys of lanthanum and nickel; the composition at maximum loading is $LaNi_5H_{6.7}$, which stores 1.37 wt% hydrogen. A very similar alloy, in which Ni is replaced partially by one or more transition metals and La may be replaced by a mixture of lanthanides, is used commercially as the cathode in NiMH rechargeable batteries. In these batteries, trivalent nickel oxide (NiOOH) is the anode. As last examples, Mg_2NiH_4 stores 3.59 wt% hydrogen, while VH_2 stores 2.1 wt% hydrogen. Other hydrides, such as MgH_2 (7.6 wt%) or AlH_3 (effective equivalent weight of 10), have been suggested, but have not been widely used so far, mainly because charging and discharging with hydrogen can be done only at an elevated temperature, and the cost of preparing these hydrides, both in terms of energy and money, is prohibitive.

Complex hydrides are inorganic materials that are best described as salts, which are built from complex anions containing hydrogen as terminal ligand, such as the BH_4^- (tetrahydroborate or borohydride) or AlH_4^- (alanate) anions and counter-cations from many different groups in the periodic table. Complex hydrides contain metallic, ionic, and covalent bonds. Clearly, the light elements and more covalent hydrides, BH_3, AlH_3, and NH_3, have high energy densities, but are very difficult to handle safely and they decompose to very stable elements, B, Al, and N_2, which are very challenging to refuel with hydrogen on board a vehicle. All three compounds readily react with ionic hydrides, forming $LiAlH_4$, $LiBH_4$, $LiNH_2$, Mg_2NiH_4, Mg_2FeH, etc. A $LiBH_4 + SiO_2$ complex hydride has been

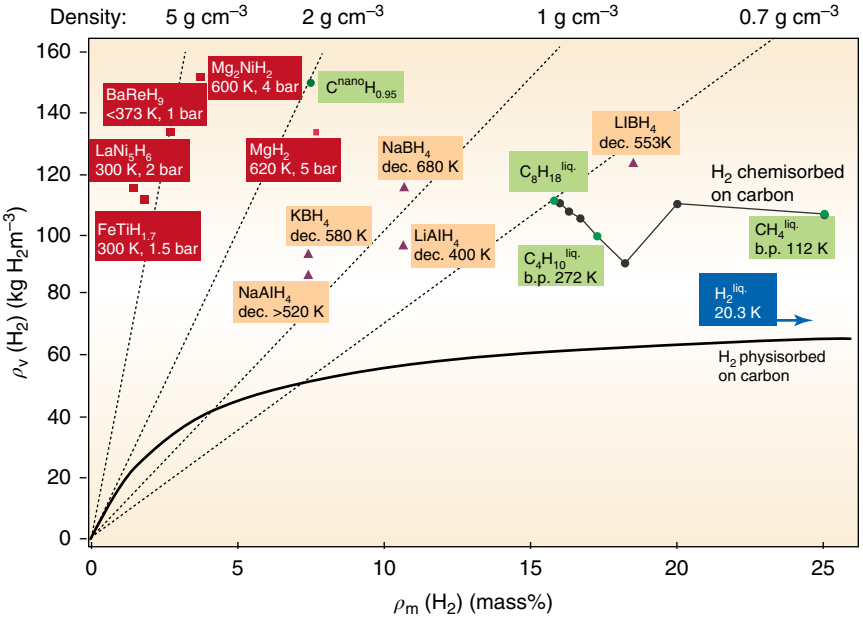


Figure 20.15 Comparison of the gravimetric and volumetric hydrogen storage of metal hydrides, carbon nanotubes, petrol, and other hydrocarbons. Interstitial (including intermetallic) hydrides are marked by red boxes, complex hydrides by orange boxes, hydrocarbons and carbon nanotubes by green boxes, and liquid hydrogen by blue box. Source: Reproduced with permission from L. Schlapbach and A. Züttel, *Nature* **414** (2001) 353.

reported to store 13.5 wt% hydrogen and release it at 60 °C. $NaAlH_4$ + porous carbon has been reported to store 7.0 wt% hydrogen, but releases it only at 400 °C. One of the highest volumetric hydrogen densities reported to date is 150 kg m^{-3}, in Mg_2FeH_6 and $Al(BH_4)_3$. For gaseous hydrogen fuel tanks to be used in vehicles, we need at least 4–5 wt%. This low gravimetric storage is the general weakness of all known metal hydrides working near room temperature (see Figure 20.15). While some hydrides can store up to 9 wt% hydrogen (e.g. $Li_3Be_2H_7$) or 4.5 H/M atomic ratio (e.g. $BaReH_9$), they are not reversible within the required range of temperature and pressure.

It should be noted that the numbers given above for total weight per unit weight of hydrogen (namely, the effective equivalent weight) do not tell the whole story. Even if hydrogen is stored chemically in the manner discussed above, it still must be kept in a suitable container, and heat exchangers must be provided to remove the heat evolved during storage and to provide the heat needed when hydrogen is discharged. Thus, the *effective equivalent weight* in an actual storage device could be significantly higher.

Attempts have been made in recent years to use nanotechnology for better hydrogen storage. CNTs and other forms of highly dispersed carbon have been tried, with only partial success. It should be realized that storage of hydrogen by bonding it chemically (as in AlH_3) or physically (as in CNTs) must satisfy two criteria that are inherently opposing each other. The thermodynamic criterion is

that bonding should be strong, so that hydrogen could be stored at, or close to, ambient temperature and pressure. On the other hand, the kinetics of releasing it on demand should be fast, and consume as little energy as possible. Now, the standard Gibbs free energy of a reaction is given by the difference between the standard Gibbs free energies of activation of the forward and reverse direction:

$$\Delta G_{\mathrm{f}}^{0\ddagger} - \Delta G_{\mathrm{b}}^{0\ddagger} = \Delta G^{0} \tag{20.50}$$

where adsorption (storage) and desorption (release) are considered to be the "forward" and "backward" reactions, respectively. A good medium for hydrogen storage should have a high negative value of ΔG^0. But, it follows from Eq. (20.50) that the standard Gibbs free energy of activation for releasing the adsorbed or absorbed hydrogen is given by

$$\Delta G_{\mathrm{b}}^{0\ddagger} = \Delta G_{\mathrm{f}}^{0\ddagger} - \Delta G^{0} \tag{20.51}$$

Considering that the energies of activation are, by definition, positive while ΔG^0 is negative, it follows that $\Delta G_{\mathrm{b}}^{0\ddagger} \geq -\Delta G^{0}$. Thus, increasing the capacity for storage, by finding a medium that strongly binds hydrogen, decreases the rate of its delivery. Catalysis cannot increase the rate of release of hydrogen because $-\Delta G^0$ in the above inequality is a thermodynamic quantity, which is not influenced by catalysis. On the other hand, increasing the surface area per unit weight could improve hydrogen storage capacity, and this is where nanotechnology may eventually lead to an effective way for hydrogen storage.

Actual and potential applications of metal hydrides include hydrogen storage, batteries, fuel cells, purification or separation of hydrogen, hydrogen getters, heat storage, heat pumps and refrigerators, thermochemical hydride compressors, sensors, detectors, and catalysts. One of the most active and interesting research on metal hydrides has been as a carrier of hydrogen fuel for motor vehicles. Hydrogen recovery is obtained in this case by simply heating the material by the exhaust gas. Increasing R&D budgets have also been invested in an effort to develop vehicles driven by fuel cells.

Further Reading

1 Hacker, V. and Mitsushima, S. (2018). *Fuel Cells and Hydrogen*. Elsevier.
2 Andwari, A.M., Pesiridis, A., Rajoo, S. et al. (2017). A review of battery electric vehicle technology and readiness levels. *Renewable Sustainable Energy Rev.* 78: 414–430.
3 Chen, G.Z. (2017). Supercapacitor and supercapattery as emerging electrochemical energy stores. *Intern. Mater. Rev.* 62: 173–202.
4 Töpler, J. and Lehmann, J. (ed.) (2016). *Hydrogen and Fuel Cell: Technologies and Market Perspectives*. Germany: Springer-Verlag.
5 Rusman, N.A.A. and Dahari, M. (2016). A review on the current progress of metal hydrides material for solid-state hydrogen storage applications. *Inter. J. Hydrogen Energy* 41: 12108–12126.
6 Sharaf, O.Z. and Orhan, M.F. (2014). An overview of fuel cell technology: fundamentals and applications. *Renewable Sustainable Energy Rev.* 32: 810–853.

7 Yu, A., Chabot, V., and Zhang, J. (2013). *Electrochemical Supercapacitors for Energy Storage and Delivery: Fundamentals and Applications.* Boca Raton, FL: CRC Press.
8 Rahman, M.A., Wang, X., and Wen, C. (2013). High energy density metal-air batteries: a review. *J. Electrochem. Soc.* 160: A1759–A1771.
9 Durbin, D.J. and Malardier-Jugroot, C. (2013). Review of hydrogen storage techniques for on board vehicle applications. *Inter. J. Hydrogen Energy* 38: 14595–14617.
10 Suh, M.P., Park, H.J., Prasad, T.K., and Lim, D.W. (2012). Hydrogen storage in metal–organic frameworks. *Chem. Rev.* 112: 782–835.
11 Armand, M. and Tarascon, J.-M. (2008). Building better batteries. *Nature* 451: 652–657.
12 Züttel, A. (2003). Materials for hydrogen storage. *Materials Today* 6: 24–33.
13 Schlapbach, L. and Züttel, A. (2001). Hydrogen-storage materials for mobile applications. *Nature* 414: 353–358.
14 Steele, B.C.H. and Heinzel, A. (2001). Materials for fuel-cell technologies. *Nature* 414: 345–352.
15 Carrette, L., Friedrich, K.A., and Stimming, U. (2001). Fuel cells – fundamentals and applications. *Fuel Cells* 1: 5–39.
16 Tarascon, J.M. and Armand, M. (2001). Issues and challenges facing rechargeable lithium batteries. *Nature* 414: 359–367.

Index

a

activated complex 53, 54, 59, 60, 67, 68, 72, 253
activated complex theory, *see* transition-state theory
activation-controlled current density 3, 35, 43, 44, 88, 159, 160, 171, 336
activation overpotential 51, 253, 396, 417, 421
activation polarization 253
activity coefficient 9, 38, 49, 101, 187, 233, 234, 242
additive manufacturing (AM) 373
additives 20, 49, 317, 318, 323, 325, 336–339, 361
admittance 198, 201, 202, 213, 221–223
adsorbed hydrogen 28, 88, 90, 91, 125, 141, 142, 152, 282, 340
adsorption isotherm 123–128, 137, 151, 192, 337, 338
adsorption pseudocapacitance 127–132, 138, 150, 190
adsorption theory 269
alkaline battery 393
alkaline fuel cell (AFC) 413, 420
alloy deposition 49, 83, 325, 328, 329, 346–350
amphoteric 248, 249, 288
angular velocity 41, 42, 128, 428
anode 2, 66, 91, 103, 105, 228, 231, 239–241, 249, 263–265, 289, 291, 293, 295–297, 308–312, 314, 315, 328–330, 334–336, 339, 351, 357, 359, 372, 374, 388, 389, 392–394, 397–399, 401–405, 408–421, 423, 432, 441
anodic current 29, 30, 185, 254, 255, 258, 266, 275, 283, 311, 355
anodic dissolution 228, 258, 266, 268, 269, 275, 291, 292, 300, 301
anodic inhibitors 313, 314, 316
anodic protection 52, 272, 303, 304, 308, 312, 313, 320
anodic reaction 29, 54, 58–60, 67, 69, 79, 142, 228, 229, 254, 256, 258, 273, 293, 314, 399
anodizing 248, 291, 318
anomalous codeposition (ACD) 349, 357
applied potential 30, 36, 49, 50, 75, 97, 117, 137, 195, 254, 274, 347, 382
AT-cut quartz crystal 214, 219
atomic force microscopy (AFM) 285, 374
atomic hydrogen 91, 141, 152, 192, 219, 220, 279, 280, 282, 284, 297, 299, 310, 326
autocatalytic 290, 292, 352
auxiliary reference electrode 20

b

batteries 52
  classes of 388–389
  quality 391–392
battery electric vehicles (BEVs) 409
battery stack 396
bimetallic corrosion, *see* galvanic corrosion
biofilms 285, 288, 305, 306
biofouling 288, 305
bi-potentiostat 45
Bode modulus plot 198, 200, 207
Bode phase-angle plot 198
Boltzmann constant 54, 280
Bond number 382, 383
bonding energy 11, 96, 137, 369
Born–Oppenheimer approximation 62–63
boundary condition 167–169, 171, 172, 186, 189, 383
Brasher–Kingsbury equation 210
breakdown potential 266, 268

Physical Electrochemistry: Fundamentals, Techniques, and Applications,
Second Edition. Noam Eliaz and Eliezer Gileadi.

Brunauer–Emmett–Teller (BET) theory 372
Butler–Volmer equation 55, 56, 252, 254, 255, 272, 277

c

capacitor 4–6, 97, 99, 104–106, 128–130, 138, 162, 197
capacity 387, 388, 398, 400, 405, 407–409, 437, 443
capillary length 382
carbon nanotubes (CNTs) 405, 442
Carnot cycle 390
catalysis 363–372, 421, 423, 443
catalyst 14, 29, 353, 368, 369, 371, 408, 414, 416, 423, 441
catalytic activity 4, 30, 83, 84, 89, 94–96, 139, 142, 282, 353, 368–370, 372
cathode 52, 61, 66, 89, 95, 231, 239, 240, 259, 264, 295, 328
cathodic inhibitors 314, 316
cathodic protection (CP) 245, 250, 263, 288, 289, 297, 303–305, 308, 310–312
cathodic transfer coefficient 68, 69
cell voltage 17, 244, 387, 393, 394, 397, 399, 424
charge transfer 1–3, 5, 30, 43, 50, 59–61, 64, 92, 123, 127, 146, 163, 166, 350, 351, 359, 417, 425
  current density 24
  resistance 5, 51, 97, 206, 255
charge-transfer current density, *see* activation controlled current density
chemical energy 1, 11, 151, 156, 239, 388–390, 412, 421
chemical potential 9, 11, 12, 31, 97, 112, 148, 233, 236, 279, 364, 388
chemical vapor deposition (CVD) 319, 320, 323
chemisorption 145, 146
chromate conversion coatings 318, 319
chronoamperometry 168, 171
chronopotentiometry 172
coating 26, 137, 204, 207, 208, 210, 222, 223, 271, 282, 283, 285, 288, 289, 291, 295, 297, 299, 303–305, 308–313, 315–320, 324, 326–329, 331, 334, 341, 342, 350, 352, 354, 357–359, 363, 368, 405
complex hydride 440, 441
complexing agent 5, 326, 337, 348, 350, 353, 354
complex-plane impedance plot 198, 202, 204–206
concentration overpotential 51
concentration polarization 204, 253, 320
convection 30, 31, 46, 378, 379
conversion coating 291, 317, 318
copper 2, 10, 15, 150, 220, 224, 231, 238, 271, 289, 327, 352, 379
corrosion
  cathodic protection 308
  current density 210, 256–258, 260, 262, 269, 274, 276, 315, 316
  fatigue 270, 303
  potential 206, 258–260, 268, 269, 272–276, 308, 315, 316
corrosion inhibitors 207, 259, 288, 299, 313, 314, 316
corrosion intensity (CI) 257
corrosion penetration rate (CPR) 257
corrosion rate (CR) 206, 241, 253, 257, 264, 277, 287, 288
coulostatic (charge-injection) method 164
counter electrode (CE) 17–19, 34, 328, 329, 331
crevice corrosion 266, 268, 289–291
critical pitting potential, *see* breakdown potential
cryogenic 436, 439, 440
crystal 12, 66, 96, 98, 137, 138–141, 149–151, 213, 214, 216–223, 253, 304, 325, 336, 340, 363, 440, 441
crystallization overpotential 52, 325
current collector 177, 390, 392, 398, 402, 405, 408, 412, 421, 425, 430
current density 2
current distribution 25, 26, 310, 328, 330, 331, 339, 348
current efficiency, *see* faradaic efficiency
current-potential relationship 58, 185
cycle life 387, 396, 400, 434
cyclic potentiodynamic polarization curves 266, 267
cyclic voltammetry (CV) 181, 183, 188–190

d

Daniell cell 238, 239
dealloying 295
Debye–Hückel theory 233
Debye length 381
deposition current density 332, 355
dezincification 295
differential-pulse polarography (DPP) 39

diffuse double layer 100, 104, 110
diffusion 3, 30, 31, 33, 35, 38, 39, 42, 86, 97, 160–163, 165–178, 184–188, 203–207, 218, 253, 262, 278, 280, 282, 283, 289, 315, 320, 325, 327, 342–344, 365, 370, 371, 379, 381, 423-425
diffusion coefficient 3, 31, 38, 41, 167, 262, 282
diffusion equation 167, 168, 172, 186
diffusion layer 3, 31, 160, 167, 170, 330, 335
diffusion limited current density 162
dip-pen nanolithography (DPN) 375
direct methanol fuel cell (DMFC) 390, 413, 415, 418
discharge rate 396, 399, 400, 433
double layer 4, 60, 99, 101, 103, 105, 107, 146, 153, 162, 236, 283, 360, 361, 430
double-layer capacitance 4, 6, 37, 97, 100, 102, 131, 145, 161, 165
double-pulse galvanostatic method 163, 165, 166
driving force 240, 296
dropping-mercury electrode (DME) 38

e

effective diffusion coefficient 283
effective equivalent weight 436, 441, 442
electric vehicle (EV) 398, 409
electrical double layer (EDL) 97, 103
electrical energy 1, 59, 60, 156, 236, 239, 323, 388–391, 417, 432
electrical potential 112, 304, 376
electroactive 3, 30, 76, 84, 87, 174, 175, 205, 356, 390
electrocapillary 113, 114, 117, 119
electrocapillary electrometer 114, 148
electrocapillary equation 112, 113
electrocatalysts 368, 369
electrochemical capacitor (EC) 388, 430, 432 *see also* supercapacitor
electrochemical cell 22, 162, 164, 215, 220, 236, 239, 280, 286, 323, 374, 389
electrochemical deposition 323, 373
electrochemical dip-pen nanolithography (E-DPN) 375
electrochemical double-layer capacitors (EDLCs) 431
electrochemical enthalpy of activation 80, 81
electrochemical fabrication (EFAB) 373
electrochemical fountain pen nanofabrication (ec-FPN) 375
electrochemical hydrogen permeation test 279, 280
electrochemical impedance spectroscopy (EIS) 84, 195, 197, 272
electrochemical jet printing 374
electrochemical nanotechnology 363
electrochemical noise (EN) 272, 278
electrochemical permeation 282, 283
electrochemical potential 9, 10, 12, 13, 28, 125, 236, 278, 279, 283, 300, 440
electrochemical printing (EcP) 363, 372, 373, 379, 380
electrochemical quartz crystal microbalance (EQCM) 28, 213, 224
electrochemistry 10, 29, 34, 37, 52, 80, 110, 114, 128, 160, 167, 195, 202, 222, 272, 331, 363, 369, 371, 380, 401, 428
electrode kinetics 1, 27, 34, 35, 38, 43, 50–52
electrode potential 105, 236
electrodeposition 46, 203, 323–326, 328, 331, 333, 338–340, 342, 344, 350, 373–376, 379
electroless deposition 352–358
electrolysis 2, 323, 373, 437, 438
electrolytic cell 30, 239
electrometer 114–119, 148
electromotive force (EMF) series 14, 236, 241
electrophoretic deposition (EPD) 323, 358
electroplating 137, 169, 221, 224, 239, 279, 297, 299, 319, 323–327, 330, 332, 335, 338, 339, 343, 346, 352, 356, 357, 359, 373, 411
electrosorption 37, 87, 121, 145–147, 150–153, 155, 157, 370, 432
electrostatic energy 11, 101, 102, 153–155
energy conversion 388, 389, 417, 419, 422
energy density 387, 406, 408, 410, 415, 418, 432, 437, 439
energy dispersive X-ray spectroscopy (EDS) 284, 379, 380
energy storage 207, 387, 388, 396, 405, 407–409, 431, 432, 434, 435
enthalpy of activation 30, 79, 81
entropy 145, 146, 155–157, 236, 280, 439
environmentally induced cracking (EIC) 297
environmental stress cracking (ESC) 227
epitaxial growth 325, 341
equilibrium constant 72, 74, 78, 123, 125, 126, 146, 151–154, 156, 235, 241

equilibrium constant (*contd.*)
equivalent circuit 6, 35, 36, 97, 98, 129, 130, 160, 161, 163, 167, 199, 201, 203, 207
equivalent weight (EW) 224, 225, 257, 390, 415, 436, 441, 442
erosion corrosion 296, 304
Evans diagram 257–265, 296, 311, 315, 316, 355, 356
excess charge density 87, 97, 99, 150
exchange current density 25, 50, 51, 55, 56, 71, 72, 77, 89, 91, 95, 252, 253, 260, 261, 263, 264, 326, 331, 332
exfoliation 294, 295
extended-range electric vehicle (E-REV) 409

f

faradaic efficiency (FE) 83, 323, 348
faradaic reaction 150, 165, 185, 254, 379
faradaic resistance 97, 199, 203, 330–332, 338
Faraday constant 3, 13, 27, 210
fast transient 20, 35, 159, 178
filiform corrosion 291
focused ion beam (FIB) 375
Fourier-transform infrared (FTIR) spectroscopy 285
fractional surface coverage 89, 96, 125, 157
frequency response analyzer (FRA) 196, 220
Frumkin isotherm 126, 129, 130, 132, 154, 191
fuel cell (FC) 67, 91, 146, 181, 184, 207, 239, 244, 338, 369, 388–390, 412, 414, 416–418, 421
  electric buses 426
fugacity 279, 280

g

galvanic cell, *see* reversible cell
galvanic corrosion 240, 277, 288, 295, 296, 308, 309, 316
galvanic couple 239, 263, 264, 273, 278, 296, 299
galvanic series 240, 296
galvanizing 319
galvanostatic 33–35, 50, 162, 163, 172
general corrosion, *see* uniform corrosion
Gibbs adsorption isotherm 111, 112, 124
Gibbs free energy 28, 29, 49, 53–55, 62, 110, 112, 123–126, 139, 146, 151, 155–157, 231, 236, 351, 364, 365, 368–370, 394, 404, 443
Gibbs free energy of adsorption 123–126, 146, 151, 155, 157
Gibbs phase rule 5
glassy metal 252
gold 96, 98, 139, 140, 142, 149, 150, 214, 220, 221, 223, 224, 227, 295, 306, 369, 384
Gouy–Chapman diffuse double layer 218
graphene 432

h

halide 231, 287, 292
hanging-dropping mercury electrode (HDME) 88
Helmholtz capacitance 105
Helmholtz double layer 105, 155
Helmholtz model 99, 103
heterogeneous catalyst 95, 368
high-temperature fuel cell (HTFC) 421–423
high-temperature PEMFC 419
Hull cell 338, 339
hybrid capacitor 432
hybrid electric vehicles (HEVs) 409
hybrid supercapacitors 432
hydration energy 324
hydride 280, 297, 341, 345, 439–442
hydrocarbon 369, 438
hydrogenation 91, 369
hydrogen attack 299
hydrogen economy 91, 436
hydrogen embrittlement (HE) 270, 297, 323, 325
hydrogen evolution reaction (HER) 14, 52, 67, 393
hydrogen-induced blistering 299
hydrogen storage 410, 436, 437, 439
hydronium 20, 77, 79, 279, 297
hydrophilic 220, 376, 383
hydrophobic 220, 376, 407

i

Ilkovic equation 38
immunity 239, 245, 249, 252, 263, 308
impedance 6, 196
impedance modulus 196
impressed-current cathodic protection (ICCP) 310
indicator electrodes 20

induced codeposition 350, 357
inductance 195
inhibitor 252, 285, 288, 289, 295, 297, 304, 313–315, 320, 338, 393
inner Helmholtz plane (IHP) 104
inner-sphere charge transfer 65
interface 1–6, 25, 27, 37, 54, 61, 68, 105, 128–130, 139, 207, 230, 269, 282, 298, 299, 303, 327, 341, 342, 357, 364–365, 382, 383, 414
intergranular corrosion 293
intermediate 53, 59, 60, 66, 69, 72, 73, 90, 94, 96, 124, 127, 146, 202, 225, 245, 247, 351, 365, 369, 370, 441
intermetallic hydride 441, 442
internal combustion engine (ICE) 390
interphacial tension 114
interphase 6, 10, 12–14, 17, 30–33, 36, 37, 51, 59–61, 63–66, 97–101, 103, 105–107, 109–117, 119, 121, 139, 145, 148, 151, 153–157, 162–167, 171, 189, 195–197, 199, 219, 220, 222, 331, 333, 346, 350–352, 354, 431, 434
interstitial hydrides 441
ion current rectification 381
ion–ion interactions 102, 124
ionic liquids 406, 407
ionic strength 101, 232–234
iron 15, 91, 229, 231, 245, 249, 250, 256, 258–260, 262–264, 286, 288, 292, 306, 309, 318, 319, 368, 441
isoelectric point 360

k

Kelvin equation 365
kinematic viscosity 40, 41

l

Langmuir isotherm 123, 125, 126, 130, 131, 152, 190, 191
Laplace operator 167
lead–acid battery 389, 397
Leclanché cell 389, 392–395, 402, 405, 411
limiting current density 3, 4, 30, 31, 41–43, 51, 159, 160, 175, 176, 223, 262, 263, 326, 342, 370
linear polarization resistance (LPR) 272, 277
linear potential sweep 36, 181, 184–186, 188, 272
linear sweep voltammetry (LSV) 181
lithium–air 408
lithium–iodine solid-state battery 395
lithium-ion battery (LIB) 402
lithium-thionyl chloride battery 393, 394
localized corrosion 231, 267, 272, 275, 277, 279, 286, 289
localized electrochemical deposition (LECD) 373
low-temperature PEMFCs 419
Luggin capillary 18, 19, 21, 23, 24, 52

m

Marcus theory 61, 63, 64
mass-transport limitation 1, 2, 6, 30, 51, 56, 77, 88, 89, 138, 159, 160, 163, 255, 330, 333, 336, 342, 347, 357, 371
mass-transport-limited current 2, 4, 31, 35, 42, 43, 84, 160, 218, 224, 262, 333
mean activity coefficient 101, 233, 234
mechanical equilibrium 115
melting point 341, 344, 345, 365, 367, 368, 422
meniscus-confined electrodeposition (MCED) 375, 376
mercury 37–39, 61, 64, 67, 69, 86, 87, 89, 92, 95, 98, 107, 111, 114–117, 119, 139, 147, 404
metal–air batteries 408
metallic coatings 317–319, 352
Mg–air batteries 408
microbiological corrosion (MIC) 227, 305
microelectrode 34, 174–178, 285, 370, 374
micro-electromechanical systems (MEMS) 324
micropipette 374, 375
migrating corrosion inhibitors 315
migration 30, 184, 360, 361, 379
mixed inhibitor 315
mixed-potential theory 252, 255, 257, 258, 263, 355
molten carbonate fuel cell (MCFC) 422
monolayer 27, 85, 86, 90, 93, 111, 123, 124, 128, 135, 137, 141, 148, 222, 383

n

Nafion 416, 417, 424
nanoelectrochemistry 379
nanoelectrode 379–381
nanoparticle 353, 363, 366, 370, 405
nanopipette 381–383
nanotechnology 363, 364, 416, 442
Nernst diffusion layer 3, 31, 35, 160, 167, 168, 174, 218

Nernst equation 5, 12, 15, 20, 49, 73, 74, 76, 78, 138, 152, 169, 186, 192, 235, 236, 241–243, 247, 250, 251, 273, 280, 348, 390
neutral species 59, 87, 89, 112, 121, 152
nickel 11, 49, 91, 106, 150, 266, 368, 399, 441
nickel–cadmium battery 398
nickel–metal hydride battery (NiMH) 401
noble metal 93, 149, 236, 273, 295, 297, 315, 327, 423
nonaqueous solution 64, 166, 343
nonpolarizable electrode 5, 49
normal hydrogen electrode (NHE) 13
normal-pulse polarography (NPP) 39
Nyquist diagram 198, 200

o

Ohmic resistance 197, 417
open-circuit potential (OCP) 255, 272, 392
organic coatings 210, 285, 303, 317
organic inhibitors 314
outer Helmholtz plane (OHP) 104, 105, 109
outer-sphere charge transfer 61, 64, 65, 351
overpotential 49–53, 55–58, 63, 64, 71, 73–75, 77, 88–91, 93, 95, 136, 163, 164, 166, 171, 186, 187, 203, 205, 238, 252, 254, 255, 262, 277, 280, 326, 331, 337, 338, 343, 347, 355, 376, 396, 408, 423
overpotential deposition (OPD) 136
oxidation reaction 15, 45, 228, 236, 238, 239, 255, 257, 282, 290, 354, 355
oxide 93, 227, 231, 248, 266, 269, 270, 288, 303, 336, 358, 364, 368, 420
oxide film theory 269
oxygen evolution reaction (OER) 91, 92
oxygen reduction reaction (ORR) 290

p

palladium 91, 220, 223, 282, 441
partial pressure 235, 241, 243, 244, 280, 287, 304
passivating inhibitor 314, 315
passivation 181, 239, 245, 252, 264, 266, 267, 269, 273, 276, 287, 289, 312, 317
passivity 239, 245, 249, 252, 264–267, 269, 270, 273
permeation flux 280, 284
permittivity 99, 102, 106, 210, 402, 428
phase shift 196, 204
phosphatizing 318
phosphoric acid fuel cell (PAFC) 412, 413
physical electrochemistry 1
physical vapor deposition (PVD) 319, 324
pickling 265, 297
pitting corrosion 52, 268, 276, 291, 316
platinum 2, 14, 20, 27, 84, 89–91, 93, 95, 106, 142, 223
plug-in hybrid electric vehicles (PHEVs) 409
poison 282, 314
polarity 64, 346, 410, 430
polarizable interface 4, 5, 37
polarization curve 255, 266–268, 273, 275, 276, 289, 302, 418
polarization resistance 56, 207, 210, 255, 256, 275, 277
polarization tests 273
polarography 37, 39, 99
polymer electrolyte membrane (PEM) 416
polymer electrolyte membrane fuel cell, *see* proton exchange membrane fuel cell
porous gas diffusion electrodes 423
potential drop 11, 18, 19, 21–24, 35, 37
potential of zero charge 87, 89, 93, 96, 99, 100, 102–104, 106, 113, 119, 138, 148–151, 153, 154, 313, 338
potential-pH diagram, *see* Pourbaix diagrams
potential sweep rate 135
potentiostat 34, 37, 45, 276–279, 282, 283, 357
potentiostatic 33–35, 37, 172, 181, 188, 276, 278, 279
Pourbaix diagrams 239, 242–245, 247–249, 252
power density 387, 391, 396, 398
pressure–composition isotherms 439
primary batteries 388, 392, 394, 397, 401
primary current distribution 329, 330, 335, 357
primary passivation potential 266
primary reference electrode 13, 19
protection potential 267, 268, 312
proton exchange membrane fuel cell (PEMFC) 413, 418
pseudocapacitance 127, 138, 150, 190, 191, 430
pseudocapacitors 432
pulse plating 341
pure capacitor 128, 197, 202

q

quartz crystal microbalance 213, 214, 216, 222
quasi-equilibrium 69, 71, 72, 79, 92, 253

r

Ragone plot 433
Randles equivalent circuit 6, 198, 199, 204
rate constant 3, 28, 55
rate-determining step (RDS) 68, 70
reaction order 74–77, 79, 89, 90, 92, 171
reaction rate 27, 34, 53, 54, 56, 60, 74, 77, 146, 253, 369
reciprocal Debye length 109
reducing agent 244, 352, 353, 355, 357, 358
reference electrode (RE) 2, 5, 18–20, 22–25
refractory metal 344
repassivation potential, *see* protection potential
resistance (Ohmic) overpotential 52
resistance polarization 253
resistor 34, 70, 129, 161, 162, 197, 199, 200, 202, 214, 277, 328, 330
resonance frequency 213, 214, 217, 218, 220, 221, 223
reversible cell 236
reversible electrode potential 236
reversible hydrogen electrode (RHE) 93, 327
reversible potential 5, 20, 49, 52, 55, 58, 72, 75, 79, 90, 93–95, 135, 137, 138, 141, 204, 242, 244–248, 258–260, 262, 273, 274, 331, 343, 348, 355, 390, 410
Reynolds number 40, 41, 382
rotating cone electrode (RConeE) 44
rotating cylinder electrode (RCylE) 272
rotating disk electrode (RDE) 40
rotating ring disc electrode (RRDE) 40, 45
roughness factor 184, 191, 218, 338, 425

s

Sand equation 172
Sauerbrey equation 216, 217
scanning electrochemical microscopy (SECM) 285, 379
scanning electron microscopy (SEM) 284, 298
scanning tunneling microscopy (STM) 139, 285
scavengers 314
secondary (rechargeable) batteries 389
secondary current distribution 329–331, 333, 335, 357
secondary ion mass spectrometry (SIMS) 285
secondary reference electrode 19
selective leaching, *see* dealloying
self-ionization 234, 244
semi-infinite linear diffusion 39, 167, 168, 174, 175, 178, 184, 205
sensitization 293, 294
separators 396, 401–403, 430, 435
silver 20, 34, 59, 140, 216, 224, 264, 306, 327
single-crystal electrochemistry 135
single-step electrode reaction 54
slippage 219, 223
solid electrolyte 93, 405, 421
solid/electrolyte interphase (SEI) 394
solid oxide fuel cells (SOFC) 359, 413, 421
solid polymer electrolyte (SPE) 416
solution resistance 6, 17, 19, 22–24, 97, 176, 188, 197, 199, 328–330, 335, 370, 371, 425, 433
solvent reorganization energy 63, 64
specific conductivity 21, 176, 371
specific energy 387, 391, 393–397, 399, 400, 402, 404, 409, 410, 412, 415, 429, 430, 433
specific power 387, 429, 433
specific resistivity 329, 330, 332
square-wave polarography (SWP) 39
standard chemical potential 13, 247, 248, 251
standard electrochemical Gibbs free energy 29, 54, 60
standard Gibbs energy 155, 191
standard hydrogen electrode (SHE) 5, 13, 14, 232, 392
standard potential 5, 14, 37, 39, 49, 74, 94, 131, 136, 138, 185, 186, 191, 236, 239, 243, 247, 251, 260, 262, 326, 327, 332, 338, 349, 350, 390, 399
static-dropping mercury electrode (SDME) 39
Stern–Geary coefficient (SGC) 256
Stern model 103–106, 360
stoichiometric coefficient 236, 242, 243
Stokes–Einstein law 282
stored charge density 99
stress corrosion cracking (SCC) 270, 300
supercapacitor 207, 388, 428–432
supercapattery 432

surface concentration 3, 5, 33, 69, 111, 128, 172, 185, 192, 280, 284, 342, 347, 366
surface coverage 69, 72, 73, 90, 155, 157, 191, 314, 340, 353
surface energy 297, 364, 368, 376, 383
surface excess 109–114, 118, 119
surface tension 37, 111–119, 364, 382, 383
sweep rate 135, 138, 181, 182, 186, 187, 189
symmetry factor 29, 54–57, 59, 64, 67, 69, 72, 79, 88, 90, 252–255, 352

t

Tafel constant 277
Tafel equation 52, 53, 61, 252, 256, 258, 272, 274
Tafel extrapolation 255, 256, 272, 275, 277
Tafel slope 39, 49, 52, 56–58, 68, 71–73, 77, 79, 81, 88–90, 92, 187, 203, 210, 256, 258, 260, 275, 315, 316, 331, 347, 355
Teflon 40, 416, 419
Temkin isotherm 127
thermodynamic equilibrium 115, 340, 383, 440
thermodynamics 1, 15, 37, 74, 79, 102, 135, 136, 145, 148, 232, 390
three-electrode setup 18
throwing power 137, 313, 324, 326, 332, 334, 343
time constant 4, 130, 162, 165, 198, 202, 206, 207, 380, 428, 429, 432, 433
titanium 91, 140, 273, 276, 305, 310, 318, 368, 441
transfer coefficient 29, 67–69, 72, 73, 77, 79, 81, 88, 90, 92, 187, 255, 347
transition-state theory 29, 53, 253, 254
two-electrode cell 6, 34, 45

u

ultramicroelectrode (UME) 174, 177, 370
uncompensated solution resistance 21–23, 36, 37, 89, 98, 163, 182, 184, 188, 189
underfilm corrosion, *see* filiform corrosion
underpotential deposition (UPD) 135, 136, 323
uniform corrosion 231, 272, 277, 286, 289

v

valve metal 93, 266
van't Hoff equation 439
vapor pressure 315, 365–368, 406, 417
vapor phase inhibitors, *see* volatile corrosion inhibitors
viscosity 40, 217–222, 282, 289, 304
volatile corrosion inhibitors (VCIs) 315
volcano plot 369
voltammogram 139, 140, 142, 151, 181, 182, 185, 189, 191, 192
volumetric density 440

w

Wagner number 329, 330, 333, 335
Warburg impedance 160, 161, 163, 166, 203, 205–207
water 28, 78, 84, 91, 104–106, 145, 146, 148, 152, 153, 155–157, 210, 222, 243–245, 248, 250, 286, 289, 294, 301, 306, 310, 358, 360, 368–370, 382, 417, 419–421, 424, 436–438
wetting 115, 116, 383, 406
work function 96, 138, 151, 152, 360
working electrode (WE) 17–22, 24, 33, 52, 80, 87, 93, 210, 220, 255, 272, 277, 279, 286, 328, 379

x

X-ray diffraction (XRD) 285
X-ray photoelectron spectroscopy (XPS) 285

y

Young–Laplace equation 114–116, 383
Young's equation 383

z

zero-resistance ammeter (ZRA) 272, 277, 278
zeta potential 105, 360
Zn–air batteries 408