modern aspects of electrochemistry no. 39

Edited by C.G. VAYENAS

MODERN ASPECTS OF ELECTROCHEMISTRY

No. 39

Modern Aspects of Electrochemistry

Modern Aspects of Electrochemistry, No. 38:

- Solid State Electrochemistry encompassing modern equilibria concepts, thermodymanics and kinetics of charge carriers in solids.
- Electron transfer processes, with special sections devoted to hydration of the proton and its heterogeneous transfer.
- Electrosorption at electrodes and its relevance to electrocatalysis and electrodeposition of metals.
- The behavior of Pt and other alloy electrocatalyst crystallites used as the electrode materials for phosphoric acid electrolyte fuel-cells.
- Applications of reflexology and electron microscopy to the materials science aspect of metal electrodes.
- Electroplating of metal matrix composites by codeposition of suspended particles, a process that has improved physical and electrochemical properties.

Modern Aspects of Electrochemistry, No. 37:

- The kinetics of electrochemical hydrogen entry into metals and alloys.
- The electrochemistry, corrosion, and hydrogen embrittlement of unalloyed titanium. This important chapter discusses pitting and galvanostatic corrosion followed by a review of hydrogen embrittlement emphasizing the formation of hydrides and their effect on titanium's mechanical properties.
- Oxidative electrochemical processes of organics, introduces an impressive model that distinguishes active (strong) and non-active (weak) anodes.
- Comprehensive discussions of fuel cells and Carnot engines; Nernst law; analytical fuel cell modeling; reversible losses and Nernst loss; and irreversible losses, multistage oxidation, and equipartition of driving forces. Includes new developments and applications of fuel cells in trigeneration systems; coal/biomass fuel cell systems; indirect carbon fuel cells; and direct carbon fuel cells.
- Exploration of the catalytic effect of trace anions in outer-sphere heterogeneous charge-transfer reactions.
- Results of the experimental and theoretical investigations on bridging electrolyte-water systems as to thermodynamic and transport properties of aqueous and organic systems. Revised version of chapter four in Number 35.

List of Contributors, MAE 39

Thomas Z. Fahidy
Department of Chemical Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

Joo-Young Go
Department of Materials Science and Engineering, Korea Advanced Institute of Science and Technology, 373-1, Guseong-dong, Yuseong-gu, Daejeon, 305-701, Republic of Korea

Jane S. Murray
Department of Chemistry, University of New Orleans, New Orleans, LA 70148

Peter Politzer
Department of Chemistry, University of New Orleans, New Orleans, LA 70148

Su-Il Pyun
Department of Materials Science and Engineering, Korea Advanced Institute of Science and Technology, 373-1, Guseong-dong, Yuseong-gu, Daejeon, 305-701, Republic of Korea

Ashok K. Vijh
Institut de recherche d'Hydro-Québec, 1800, Blvd. Lionel-Boulet, Varennes, Québec Canada J3X 1S1

Gregory X. Zhang
Teck Cominco, Product Technology Centre, Mississauga, Ontario, Canada

MODERN ASPECTS OF ELECTROCHEMISTRY

No. 39

Edited by

C. G. VAYENAS
University of Patras
Patras, Greece

RALPH E. WHITE
University of South Carolina
Columbia, South Carolina

and

MARIA E. GAMBOA-ADELCO
Managing Editor
Superior, Colorado

Springer

C. G. Vayenas
Department of Chemical Engineering
University of Patras
Patras 265 00
Greece
cat@chemeng.upatras.gr

Library of Congress Control Number: 2005939188

ISBN-10: 0-387-23371-7 e-ISBN 0-387-31701-5
ISBN-13: 978-0387-23371-0

Printed on acid-free paper.

Printed in the United States of America. (TB/IBT)

9 8 7 6 5 4 3 2 1

springer.com

Volume 39 is dedicated to the memory of Professor Brian E. Conway, a leading electrochemist, a splendid teacher, a wonderful friend, a great man, whom the entire electrochemical community will always remember.

John O'M Bockris,
Ralph White,
Costas Vayenas, and
Maria Gamboa-Aldeco

Preface

This volume of Modern Aspects covers a wide spread of topics presented in an authoritative, informative and instructive manner by some internationally renowned specialists. Professors Politzer and Dr. Murray provide a comprehensive description of the various theoretical treatments of solute-solvent interactions, including ion-solvent interactions. Both continuum and discrete molecular models for the solvent molecules are discussed, including Monte Carlo and molecular dynamics simulations. The advantages and drawbacks of the resulting models and computational approaches are discussed and the impressive progress made in predicting the properties of molecular and ionic solutions is surveyed.

The fundamental and applied electrochemistry of the silicon/electrolyte interface is presented in an authoritative review by Dr. Gregory Zhang, with emphasis in the preparation of porous silicon, a material of significant technological interest, via anodic dissolution of monocrystalline Si. The chapter shows eloquently how fundamental electrokinetic principles can be utilized to obtain the desired product morphology.

Markov chains theory provides a powerful tool for modeling several important processes in electrochemistry and electrochemical engineering, including electrode kinetics, anodic deposit formation and deposit dissolution processes, electrolyzer and electrochemical reactors performance and even reliability of warning devices and repair of failed cells. The way this can be done using the elegant Markov chains theory is described in lucid manner by Professor Thomas Fahidy in a concise chapter which gives to the reader only the absolutely necessary mathematics and is rich in practical examples.

Electrochemical processes usually take place on rough surfaces and interfaces and the use of fractal theory to describe and characterize the geometric characteristics of surfaces and interfaces can be of significant importance in electrochemical process description and optimization. Drs. Joo-Young Go and

Su-Il Pyun provide a comprehensive review of the physical and electrochemical methods used for the determination of surface fractal dimensions and of the implications of fractal geometry in the description of several important electrochemical systems, including corroding surfaces as well as porous and composite electrodes.

Electrochemistry plays an increasingly important role in biology and medicine and the electrochemical treatment of tumors (ECT) is receiving considerable attention as a viable alternative to the more classical tumor treating approaches of surgery and chemotherapy. Dr. A. Vijh, a specialist in this area, describes both the phenomenology and the proposed physicochemical mechanisms of ECT in a comprehensive chapter.

C.G. Vayenas

R.E. White

MEMORIES OF BRIAN EVANS CONWAY
EDITOR 1955-2005

By John O'M. Bockris (Editor 1953-2003)

I remember meeting Brian Conway on the road, Queens Gate, which leads from Imperial College, London University, to the Tube Station in South Kensington, London. It was a Saturday afternoon in 1946. I was one year into a lectureship (assistant professorship in U.S. terms) at Imperial College, the technologically oriented part of London University. Brian stopped me to say that he wanted to do his graduate research with me. He was 20 and I, 23.

Brian Conway entered my research group at Imperial College in its second year. There are some who have called the years between 1946 and 1950 as the Golden Years of Electrochemistry because the development of the subject as a part of physical chemistry was pioneered in those years, primarily by the remarkable set of people who were in the group at that time and who spread out with their own students and ideas after they left me.

First among those in respect to the honors achieved was Roger Parsons. He became an FRS and President of the Faraday Division of the Chemical Society. Conway is No. 2 in this formal ranking although he published more extensively and with more original ideas than did Parsons. Then there was John Tomlinson who became a professor at the University of Wellington in New Zealand (and Vice Chancelor of the University). E.C. Potter too was in the group at that time and became in the 1980's the President of the Royal Society of New South Wales, apart from having an active career in the Council of Science and Industry of the Australian government. Harold Eagan was there too and he later achieved a grand title: "The Government Chemist," the man in the U.K. government who is in charge of testing the purity of various elements of the country's supply.

The only woman in the group, Hanna Rosenburg, should be mentioned here because she had an extended influence on Brian Conway, based partly on her remarkable ability to discuss widely.

Martin Fleischmann, too, is relevant. He got his Ph.D. in a small group near to my group in which Conway worked. However, he joined us in various activities, particularly the discussions. He became well known internationally not only because of his contributions to physical electrochemistry, but also because in 1989 he resuscitated an idea, - which had been introduced by the French and Japanese in the 1960's, - that nuclear reactions could be carried out in solutions in the cold.

Conway's thesis (1949) was very original in content and contained, e.g., 3 D representations of potential energy surfaces.

After Brian Conway got his Ph.D. with me, he went to work with J.A.V. Butler at the Chester-Beatty Cancer Research Institute, a short walk from Imperial College. Butler's habit of pouring over manuscripts, murmuring, and his sensitivity to noise whilst thinking, was reflected by some of Conway's habits. Butler was known to be absent minded in a way which colored the great respect with which he was regarded. Thus, according to Conway, in Butler's year at Uppsala he was observed by a British colleague to pass two lumps of sugar to the cashier whilst thoughtfully adding the ore to his coffee.

Contact with Conway continued from 1949 to 1953 in London. We used to meet on Saturday afternoons in a Kensington coffee shop and it was there that some of our electrochemical problems were worked out.

After 1951, I decided to look towards the United States for continuation of my career. We thought of Brian as a possible post doc to take with us and wanted to get him married before the transition. Unfortunately, Brian evaluated girls principally by their intellectual abilities and the Viennese Ms. Rosenburg had given him the model. He found it difficult to meet her equal and we had eventually to point out to him that he couldn't get a Visa to come to America with a woman unless he were married to her. Shortly, after this, he did invite me to Daguize, our Saturday café, to meet a former Latvian pharmacist, Nina, and later I served to take Nina up the aisle at the wedding of Nina and Brian (1954).

It had seemed to me that, if we were going to bring Electrochemistry from its moribund state of the 1940's to modernity in physical chemistry, a yearly monograph would help and I therefore approached Butterworth's in London with a proposition and the first volume was published in 1954.[1] I brought Brian into the Aspects in Volume I as an assistant and also my co-author in a chapter on solvation. It soon seemed unfair to continue to receive Brian's help in editing unless he was made an editor too and therefore, from the second volume on, we were both named editors of the series and kept it that way until, in 1987, we invited Ralph White to join us as the Electrochemical Engineer, thus increasing the breadth of the topics received.[2]

I moved to the University of Pennsylvania in Philadelphia in 1953 whilst Brian remained at Chester-Beatty, but I went back to London for the summer of 1954 and persuaded Brian to come over to Philadelphia at first as a post doctoral fellow.

Brian Conway's move to the U.S.A. facilitated our joint editorship in running the series. We had had many discussions in London and this encouraged us to make the composition of each volume an occasion for a Meeting, lasting two to three days in a resort. This meant one to two day's hard discussion deciding the topics we wanted and then the authors who might write the articles, together with a reserve author in case the first one refused. The third day we spent as free time for our own discussions. Colorado Springs was a place in which we had an interesting time and then there was the Grand Hotel in White Sulfur Springs and several resorts in Canada. But there is no doubt that Bermuda was more often the locale of the Modern Aspects invitation meetings than any other. After Ralph White threw his expertise onto choice of authors, we continued to meet yearly, e.g., in Quebec City and in San Francisco, though in recent years the meetings drifted back to their original form.

---

[1] Butterworth's was purchased by Plenum Press in 1955 after the second volume for Modern Aspects of Electrochemistry had been published.

[2] Our policy throughout, - as I believe it is with the present editors, - was to make in Modern Aspects of Electrochemistry a truly broad presentation of the field although we always avoided analytical topics for which there was already so much presentation.

Coming now to Brian's two years at the University of Pennsylvania (1954-1955), we got out of this time three papers which I think of as being foundation papers of what we now know as physical electrochemistry. The first concerned the mechanism of the high mobility of the proton in water. The second arose out of the interpretation of plots of log i_o against the strength of the metal-hydrogen bond and found two groups of metals, one involving proton discharge onto planar sites and the other proton discharge (rate determining) onto adsorbed H. The last of the three concerned the early stages of metal deposition, a subject which raises many fundamental questions in electrochemical theory (e.g., partial charge transfer). The first two of these works were published in the Journal of Chemical Physics and the last one in the Proceedings of the Royal Society. To have papers accepted in these prestigious journals is unusual and particularly for electrochemical topics so often mired in lowly technological issues.

Conway transferred to the University of Ottawa in 1955 and remained there for the rest of his career. He published near to 400 papers whilst in Ottawa and it's interesting to mention a few of them.

In Ottawa I would pick out the work with Currie in 1978 on pressure dependence of electrode reactions. It gave a new mechanism criterion.

Agar had suggested in 1947 that there might be a temperature dependence of β, the electrode kinetic parameter, and Conway took this up in 1982 and showed experimentally that in certain reactions this was the case.

The most well known work that Conway and his colleagues completed in Ottawa was on the analysis of potential sweep curves. I had been critical of the application of potential sweep theory to reactions which involved intermediates on the electrode surface and, working particularly early with Gilaedi and then with Halina Kozlowska, and to some extent with Paul Stonehart, Conway developed an analysis of the effect of intermediate radicals on the shape and properties of potential sweep showing how interesting electrode kinetic parameters could be thereby obtained.

This is the point to stress the part played by Halina Angerstein-Kozlowska in Brian's work. She played an important role in administration of the co-workers, apart from the supervision

of the work, particularly in Brian's absences at meetings, many in Europe. Her position was similar to that of the late Srinivasan in respect to John Appleby electrochemical research group at Texas A&M. It is more creative and directional than her formal position as a senior researcher indicates.

I continued an active discussion relationship with Brian Conway for more than fifty years. This was made possible not only by the yearly meetings for Modern Aspects but by letters, most which were in discussions of topics apart from electrochemistry. I remember those on the inheritance of acquired characteristics with Hanna Rosenburg in London in the fifties and the discussions of the position and number of the galaxies and then later, consideration of methods by aggressive groups could be restrained from war. In more recent years we have been concerned with the speed of electrons in telephone lines. Many people think that electrons there must travel very quickly but in fact their movement during the passage of messages is measured in a few cms per second.

Although the correspondence with me concerned topics discussed in terms of the science of the day, Conway's interests were broad and he had interests in the Arts, stimulated by his wife in this direction (his house was decorated with her paintings). In this respect, Conway exemplifies the Scientist as a Great Man, a man of Knowledge in all and any direction. One is reminded of Frumkin in Moscow who kept abreast of modern literature from Anglo-America, France and Germany and Hinshelwood at Oxford, a physical chemist whose alternative interests were in Japanese Poetry.

The long lasting scientific correspondence with Conway started to ebb in the mid nineties and I assumed that this was due to normal processes of aging but it turned out that it was the beginning of the end, - he was frequently in the hospital.

One day in this Summer of 2005, Nina called to tell me that the doctors declared the prostate cancer which he had suffered had metastasized and that they could no longer control it. I placed myself at Brian's disposal in respect to anything that he wanted, - I worked here through the well-known Barry Macdougall, an ex Conway student. Through him I passed to Brian photographs from the London days, and these contained, among others, the one he most

wanted, - that of the student he had known 55 years earlier, Hanna Rosenburg.

Brian leaves behind him Nina (later a professional artist and musician) who had contributed so much to his career by her help and fortitude and a son, Adrian, who is a competent theorist in telephone networks.

Brian chose to have his body cremated.

JOMB/ts

Contents

Chapter 1

QUANTITATIVE APPROACHES TO
SOLUTE-SOLVENT INTERACTIONS

Peter Politzer and Jane S. Murray

I. Introduction 1

II. General Interaction Properties Function (GIPF) 3

1. Background 3
2. Applications 9
3. Summary 16

III. Procedures that Directly Address the Solute-Solvent Interaction 17

1. Discrete Molecular Models of Solvent 17
2. Continuum Models of Solvent 26
3. Hybrid/Intermediate Models of Solvent 38
4. General Comments: Discrete Molecular and Continuum Solvent Models 41
5. Ionic Solvation 41

IV. Linear Solvation Energy Relationships (LSER) 51

1. Empirical 51
2. Theoretical 53

V. Discussion and Summary 55

Acknowledgement 56
References 56

Chapter 2

POROUS SILICON: MORPHOLOGY AND FORMATION MECHANISMS

Gregory X. Zhang

I. Introduction 65
II. Formation of Porous Silicon 71
1. I-V curves and formation condition 71
2. Hydrogen evolution and effective dissolution valence 75
3. Growth rate of porous silicon 76
4. Chemical dissolution during PS formation 78
III. Morphology of PS 79
1. Pore diameter 79
2. Pore orientation and shape 87
3. Pore branching 89
4. Interface between PS and silicon 91
5. Depth variation 92
6. Summary 96
IV. Anodic reaction kinetics 98
1. Potential distribution and rate limiting process 98
2. Reaction paths 101
3. Effect of geometric elements 103
4. Potential drop in the substrate 107
5. Surface lattice structure 110
6. Reactions on the surfaces of silicon and silicon oxide 112
7. Distribution of reactions and their rates on pore bottoms 115
8. Formation mechanisms of morphological features 118
V. Summary 128
References 130

Chapter 3

MODELING ELECTROCHEMICAL PHENOMENA VIA MARKOV CHAINS AND PROCESSES

Thomas Z. Fahidy

I. Introduction 135
II. A concise theory of Markov chains and processes 137
1. Fundamental Notions of Markov chains 137
2. The Erlang Formula 140
3. Service System with Exponential Holding Time 141
III. Markovian modeling of electrode surface processes 142
1. Random Walk of an Ion Between Adjacent Active Centres 142
2. Markovian Analysis of Layer Thickness on a Surface 143
3. Markovian Interpretation of an Anodic Deposit Formation ↔ Deposit Dissolution Process 145
4. The Electrode Surface as a Multiple Client – Server: A Markovian View 147
IV. Markovian Modeling of Electrolyzer (Electrochemical Reactor) Performance 148
1. The Batch electrolyzer 148
2. The tank electrolyzer – mixer cascade 150
3. Markovian Analysis of the CSTER 152
V. The Markov – Chain approach to miscellaneous aspects of electrochemical technology 153
1. Repair of Failed Cells 153
2. Analysis of Switching – Circuit Operations 157
3. Reliability of Warning Devices 158

4. Monitoring of Parasitic Reaction Sequence in an Electrolyzer 159
5. Flow Circulation in an Electrolytic Cell 160
VI. Final Remarks 161
Acknowledgments 164
Appendix 164
References 165

Chapter 4

FRACTAL APPROACH TO ROUGH SURFACES AND INTERFACES IN ELECTROCHEMISTRY

Joo-Young Go and Su-Il Pyun

I. Introduction 167
II. Overview of fractal geometry 170
1. Self-Similar Fractal 171
2. Self-Affine Fractal 172
III. Characterization of rough surfaces and interfaces based upon fractal geometry: methods needed for the determination of the surface fractal dimension 174
1. Physical Methods 174
2. Electrochemical Methods 185
IV. Investigation of diffusion towards self-affine fractal interface 192
V. Application of the fractal geometry in electrochemical system 209
1. Corroded Surface 209
2. Partially Blocked Active Electrode and Active Islands on Inactive Electrode 213
3. Porous Electrode 218
VI. Concluding Remarks 219
Acknowledgements 220
Notation 220
References 225

Chapter 5

PHENOMENOLOGY AND MECHANISMS OF ELECTROCHEMICAL TREATMENT (ECT) OF TUMORS

Ashok K. Vijh

I. Introduction and Origins 231
II. Phenomenology of ECT 234
III. Mechanistic Aspects 240
1. General 240
2. Electrochemical Treatment (ECT) as a case of electroosmotic dewatering (EOD) of the cancer tissue and related effects 241
3. Some possible factors in the necrosis of cancer tumors by ECT 248
4. Approaches to the possible improvement of the electrochemical treatment (ECT) of cancer tumors 249
5. In-vitro studies on the suppression of cell proliferation by the passage of electric current 251
IV. Electrochemical treatment (ECT) of tumors in animals 254
1. Mice 254
2. Hamsters 256
3. Rats 256
4. Pigs, Rabbits and dogs 260
V. Electrochemical treatment (ECT) of tumors in Human Beings: Clinical studies on patients 261
1. The new Era in ECT 262
2. Extensive clinical studies in China 263
3. Clinical studies in other countries 268
VI. Conclusions 269
References 270
Index 275

1

Quantitative Approaches to Solute-Solvent Interactions

Peter Politzer and Jane S. Murray

Department of Chemistry, University of New Orleans, New Orleans, LA 70148

I. INTRODUCTION

The interactions between solutes and solvents are noncovalent in nature (barring the occurrence of chemical reactions), and therefore fall into the same category as those that govern molecular recognition processes, the formation and properties of liquids and solids, physical adsorption, etc. Hydrogen bonding, in its many manifestations, is a particularly prominent and important example.

It follows from the Hellmann-Feynman theorem[1,2] that these interactions are essentially Coulombic, and hence can be described accurately in terms of classical electrostatics, given the correct charge distributions. This has been discussed by Hirschfelder et al.[3,4] The problem, with regard to solute-solvent interactions, is one of scale, i.e., the large number of molecules of the solvent that must be taken into account in order to realistically represent its effects. Thus, whereas interaction energies of relatively small noncovalently-bound complexes are

Modern Aspects of Electrochemistry, Number 39, edited by C. Vayenas et al., Springer, New York, 2005

increasingly being computed reliably, either directly, by Eq. (1),[5,6]

$$\Delta E = E_{complex} - \sum E_{reactant} \qquad (1)$$

or by sophisticated perturbation methods[7,8] such as symmetry-adapted perturbation theory (SAPT), the situation is different for solute-solvent systems. The direct approach, Eq. (1), is impractical; perturbation theory can be applied, but at a lower level, in establishing the intermolecular potentials for molecular dynamics and Monte Carlo simulations of solutions.[8-11]

Quantitative models of solute-solvent systems are often divided into two broad classes, depending upon whether the solvent is treated as being composed of discrete molecules or as a continuum. Molecular dynamics and Monte Carlo simulations are examples of the former;[8-11] the interaction of a solute molecule with each of hundreds or sometimes even thousands of solvent molecules is explicitly taken into account, over a lengthy series of steps. This clearly puts a considerable demand upon computer resources. The different continuum models,[11-16] which have evolved from the work of Born,[17] Bell,[18] Kirkwood,[19] and Onsager[20] in the pre-computer era, view the solvent as a continuous, polarizable isotropic medium in which the solute molecule is contained within a cavity. The division into discrete and continuum models is of course not a rigorous one; there are many variants that combine elements of both. For example, the solute molecule might be surrounded by a first solvation shell with the constituents of which it interacts explicitly, while beyond this is the continuum solvent.[16]

An overview of these methods will be presented later in this Chapter. They will be introduced largely in the context of molecular solutes, although they can also be applied to ionic ones, as shall be discussed in a separate Section III.5. First, however, shall be summarized in some detail an approach that is less comprehensive than those that have been mentioned, but is readily applied and can be quite effective.

II. GENERAL INTERACTION PROPERTIES FUNCTION (GIPF)

1. Background

Since noncovalent interactions are Coulombic, it seems reasonable to assume that a dominant role is played by the electrostatic potentials on the molecules' surfaces. This is what they *see* and *feel* as they approach each other. Qualitatively, it would be anticipated that interaction would be promoted by a complementary orientation of these potentials, i.e., positive regions on A being in the proximity of negative ones on B. (Already in 1925, Langmuir drew attention to the importance of molecular surface *fields of force* in solution interactions.)[21] Furthermore, it seems possible that if one partner in a given type of process (e.g., solvation) is held constant, then the strength of the interaction might correlate with the features of the electrostatic potentials on the molecular surfaces of its various partners. For example, perhaps the solubilities of a series of solutes in a given solvent could be related to their molecular surface electrostatic potentials.

In quantifying these ideas, certain questions must be addressed. First, how should a molecular surface be determined? It has no rigorous basis, but being a useful concept, a number of approaches have been suggested,[22-26] such as defining it in terms of a set of intersecting spheres centered at the nuclei and having some appropriate radii, e.g., van der Waals. We choose instead to identify the surface with an outer contour of the molecular electronic density $\rho(\mathbf{r})$.[25,26] This has the advantage that the surface reflects the specific features of the particular molecule, for instance lone pairs and strained bonds. Following Bader et al.,[26] we use $\rho(\mathbf{r}) = 0.001$ au (electrons/bohr^3), which they showed (for a group of hydrocarbons) to encompass more than 97% of the total electronic charge.

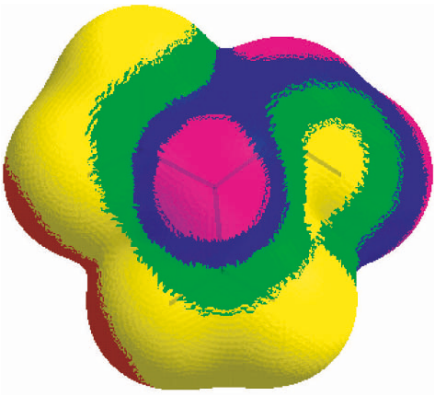
Before proceeding, it is necessary to present some general background concerning the molecular electrostatic potential. Its value $V(\mathbf{r})$ at any point $\mathbf{r}$ in the space around a molecule is the resultant of the positive contribution of the nuclei and the negative one of the electrons:

$$V(\mathbf{r}) = \sum_{A} \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}|} - \int \frac{\rho(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r}' - \mathbf{r}|} \quad (2)$$

In Eq. (2), Z_A is the charge on nucleus A, located at $\mathbf{R}_A$, and $\rho(\mathbf{r})$ is the molecule's electronic density. Thus whether $V(\mathbf{r})$ is positive or negative in any given region depends upon whether the effect of the nuclei or the electrons is dominant there. In our experience, negative $V(\mathbf{r})$ are usually associated with (a) the lone pairs of electronegative atoms, e.g., N, O, F, Cl, etc., (b) the π electrons of unsaturated hydrocarbons, such as arenes and olefins, and (c) strained C–C bonds, as in cyclopropane. It should be noted that $V(\mathbf{r})$ is a real physical property, an observable; it can be determined experimentally, by diffraction techniques,[27,28] as well as computationally.[29] Our present interest is in the electrostatic potential on the molecular surface, which will be designated $V_S(\mathbf{r})$. It is customary to report $V(\mathbf{r})$ and $V_S(\mathbf{r})$ in units of energy rather than potential, e.g., kcal/mole or kJ/mole. The values correspond, therefore, to the interaction energy between the electrostatic potential of the molecule at the point $\mathbf{r}$ and a positive unit point charge situated at that point. An example of a molecular surface electrostatic potential is shown in Figure 1, for the dimethylnitrosamine molecule, $(CH_3)_2N–NO$.

A likely application of $V_S(\mathbf{r})$ that comes immediately to mind is in regard to hydrogen bonding. It seems reasonable that the most negative (or minimum) value of $V_S(\mathbf{r})$, labeled $V_{S,min}$, should indicate the preferred site for accepting a hydrogen bond, while the most positive, $V_{S,max}$, when associated with a hydrogen should reflect propensity for hydrogen-bond donation. This has been found to be the case, as shall be discussed in Section II.2.ii. (It is certainly possible for a molecule to have two or more local surface minima and/or maxima, and thus potential hydrogen bonding sites; examples are the DNA bases.)[30]

The extrema of $V_S(\mathbf{r})$ are, however, only the beginning of the useful information that can be gleaned from it. The question is how to characterize the key features and overall pattern of $V_S(\mathbf{r})$ sufficiently to permit quantitative correlations with physical properties. Over a period of several years, we have identified a group of statistically-defined quantities that are effective for this



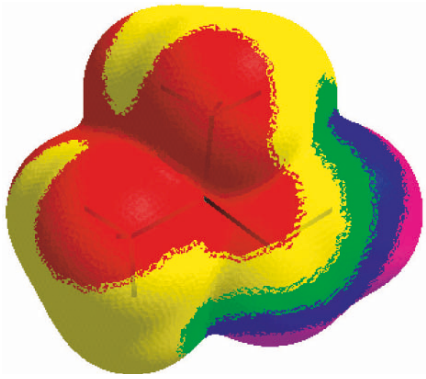


Figure 1. Electrostatic potential on the molecular surface of dimethylnitrosamine, computed at the Hartree-Fock STO-5G*//STO-3G* level. Two views are presented; the top shows the lone pair (purple color) on the amino nitrogen. Color ranges, in kcal/mole, are: purple, more negative than -20; blue, between -20 and -10; green, between -10 and 0; yellow, between 0 and +13; red, more positive than +13.

purpose. (For reviews, see Murray and Politzer.)[31-34] Obvious ones after $V_{S,max}$ and $V_{S,min}$ are the positive, negative and overall averages of $V_S(\mathbf{r})$:

$$\overline{V}_S^+ = \frac{1}{\alpha}\sum_{i=1}^{\alpha} V_S^+(\mathbf{r}_i) \tag{3}$$

$$\overline{V}_S^- = \frac{1}{\beta}\sum_{j=1}^{\beta} V_S^-(\mathbf{r}_j) \tag{4}$$

$$\overline{V}_S = \frac{1}{n}\left[\sum_{i=1}^{\alpha} V_S^+(\mathbf{r}_i) + \sum_{j=1}^{\beta} V_S^-(\mathbf{r}_j)\right] \tag{5}$$

In Eqs. (3) and (4), the summations are over the points on the positive and negative portions of the surface, respectively; n is the total number of points, including those where $V_S(\mathbf{r}) = 0$. On most molecular surfaces, the positive regions are larger in area but the negative ones are stronger;[33,35,36] thus $\left|\overline{V}_S^-\right| > \overline{V}_S^+$. Nevertheless, $\overline{V}_S$ (the overall average) is usually weakly positive, reflecting the greater extent of $\overline{V}_S^+(\mathbf{r})$. This pattern can change, however, when the molecule has several strongly-electron-withdrawing components, meaning that each receives less of the polarizable electronic charge;[33,36] this causes $\left|\overline{V}_S^-\right|$ to diminish and may lead to $\left|\overline{V}_S^-\right| < \overline{V}_S^+$. Thus, whereas CH_3F has $\overline{V}_S^+ = 8.2$ and $\overline{V}_S^- = -14.7$ kcal/mole, CF_4 has $\overline{V}_S^+ = 11.5$ and $\overline{V}_S^- = -4.6$ kcal/mole,[36] (Hartree-Fock STO-5G*//STO-3G* calculations).

In addition to these basic quantities, there are three that provide more detailed information:

1. The average deviation, Π,

$$\Pi = \frac{1}{n}\sum_{k=1}^{n}\left|V_S(r_k) - \overline{V}_S\right| \tag{6}$$

is a measure of the internal charge separation that exists even in molecules with zero dipole moment, e.g., *para*-dinitrobenzene. Π varies in magnitude from 2-5 kcal/mole for hydrocarbons to mid-20's for polynitro/polyaza systems;[33,35,36] it is 16.5 kcal/mole for *para*-dinitrobenzene. Π correlates with various empirical indices of polarity.[35,37]

2. The positive, negative and total variances, σ_+^2, σ_-^2 and σ_{tot}^2:

$$\sigma_{tot}^2 = \sigma_+^2 + \sigma_-^2 = \frac{1}{\alpha}\sum_{i=1}^{\alpha}\left[V_S^+(r_i) - \overline{V}_S^+\right]^2 + \frac{1}{\beta}\sum_{j=1}^{\beta}\left[V_S^-(r_j) - \overline{V}_S^-\right]^2 \tag{7}$$

The variances indicate the variabilities, or ranges, of the positive, the negative and the overall electrostatic potentials on the molecular surfaces. They are particularly sensitive to the extrema, $V_{S,max}$ and $V_{S,min}$, due to the terms being squared. This also means that they may be much larger than Π, by as much as an order of magnitude, and show quite different trends.[33,36,37] Basically, σ_+^2 and σ_-^2 help to show how strong are the positive and negative potentials.

3. We have introduced a quantity, ν, that is intended to reflect the degree of balance between σ_+^2 and σ_-^2, and thus to indicate to what extent a molecule is able to interact with others (whether strongly or weakly) through both its positive and negative regions.

$$\nu = \frac{\sigma_+^2 \sigma_-^2}{\left[\sigma_{tot}^2\right]^2} \tag{8}$$

From Eq. (8), it can be seen that the upper limit of ν is 0.250, which it reaches when $\sigma_+^2 = \sigma_-^2$.

The quantities that have been presented do effectively characterize the electrostatic potential on a molecular surface. We have shown that a number of macroscopic, condensed-phase properties that depend upon noncovalent interactions can be expressed in terms of some subset of these quantities (frequently augmented by the positive, negative or total surface areas, A_S^+, A_S^- or A_S). It is necessary to first establish a reliable experimental database for the property of interest, and then to fit it, by means of a statistical analysis code, to (usually) three or four of the quantities, appropriately selected, as computed for the molecules in the database. If the interaction involves multicomponent systems, as does solvation, then only one component may vary. For example, a relationship could be developed for a series of solutes in a particular solvent, or a given solute in different solvents. In doing so, we have always sought to use as few of the computed quantities as is consistent with a good correlation, since they can provide insight into the physical factors that are involved in the interaction; this becomes obscured if many terms are involved.

Once a relationship has been established for a property of interest, for instance the heat of vaporization, then the latter can be predicted for any compound, even one not yet synthesized, by computing the relevant surface quantities for a single molecule of it. The effect of the surroundings is implicitly taken into account through the correlation that has been developed.

This is accordingly a unified approach to representing and predicting condensed phase properties that are determined by noncovalent interactions. We summarize it conceptually as a *general interaction properties function* (GIPF), Eq. (9):

$$\text{Property} = f\Big[A_S^+, A_S^-, A_S, \overline{V}_S^+, \overline{V}_S^-, \overline{V}_S, \Pi, \sigma_+^2, \sigma_-^2, \sigma_{tot}^2, \nu, V_{S,max}, V_{S,min} \Big] \quad (9)$$

$V_{S,max}$ and $V_{S,min}$ are site-specific, in that they refer to a particular point on the surface. (In rare instances, we have also used V_{min}, the overall most negative value of the electrostatic potential in the three-dimensional space of the molecule; V_{min} is also site-specific.) The remaining quantities in Eq. (9) are termed *global*, since they reflect either all or an important portion of the molecular surface. It should be emphasized that on no occasion have we used more than six computed quantities in representing a property; three or four is typical. It should also be noted that the specific value of $\rho(\mathbf{r})$ chosen to define the molecular surface is not critical, as long as it corresponds to an outer contour; we have shown that $\rho(\mathbf{r})$ = 0.0015 or 0.002 au would be equally effective.[35,38] (The numerical coefficients of the computed quantities would of course be somewhat different, but the correlation would not be significantly affected.)

The GIPF technique has been used to establish quantitative representations of more than 20 liquid, solid and solution properties,[31-34] including boiling points and critical constants, heats of phase transitions, surface tensions, enzyme inhibition, liquid and solid densities, etc. Our focus here shall be only upon those that involve solute-solvent interactions.

2. Applications

(*i*) *General Comments*

Some of the properties for which GIPF expressions have been developed reflect only interactions between molecules of the same compound; examples are boiling points, critical constants, heats of phase transitions, etc. In such instances, σ_{tot}^2 and ν are often important quantities because the interactions can be strongly attractive only if both the positive and the negative surface potentials achieve relatively large magnitudes; this implies a high

value of σ^2_{tot}, and ν being near 0.250. In fact some of these properties can be represented solely in terms of the product $\nu\,\sigma^2_{tot}$ and the area.[31-34]

In solute-solvent systems, the situation can be quite different; for example, a strongly positive region on one component and a strongly negative one on the other can suffice to produce a high solubility. Such cases will be discussed in some detail. Before proceeding to this, however, two practical points should be addressed. First, how dependent are the computed surface potential quantities upon molecular conformation (since this may change in solution)?[39] This has been investigated,[40] and found to usually not be an important issue for most of the quantities, provided that some intramolecular interaction, such as hydrogen bonding, is not significantly affected; σ^2_+ is most likely to change. An example in which conformation perhaps does play a key role will be mentioned in Section II.2.iv. Second, how sensitive is $V_S(\mathbf{r})$, and hence the GIPF procedure, to the computational level at which it is obtained? This has also been studied.[41,42] and the results indicate that satisfactory GIPF correlations can generally be obtained at different levels, although the functional forms may differ somewhat. Since the molecules are sometimes relatively large, we have normally used Hartree-Fock STO-5G* $V_S(\mathbf{r})$, although in some instances we have gone to larger basis sets.

(*ii*) *Hydrogen Bonding*

There is a long history of relating hydrogen bonding to molecular electrostatic potentials; the early work has been summarized by Politzer and Daiker.[43] We have shown that hydrogen-bond-donating and -accepting tendencies, as measured by experimentally-based solvatochromic parameters, can be expressed linearly in terms of just $V_{S,max}$ and $V_{S,min}$, respectively;[44,45] no other surface quantities are needed. The former relationships are excellent, with correlation coefficients $R \geq 0.97$; the latter are less so, $R = 0.90$, perhaps because different acceptor heteroatoms are involved, whereas it is always a hydrogen that is donated.

(*iii*) *Free Energies of Solvation*

The free energy of solvation is a key property of solute-solvent systems, since it quantifies the tendency of the solute to enter into solution. We have addressed this for aqueous solvation via the GIPF approach on several occasions.[46-48] Our best correlation is Eq. (10), obtained using density functional B3P86/6-31+G** $V_S(\mathbf{r})$:[46]

$$\Delta G_{solvation} = \alpha V_{min} - \beta(V_{S,max} - V_{S,min})^3 + \gamma A_S^- \overline{V}_S^- + \delta(A_S^- \overline{V}_S^-)^{-1} + \varepsilon \quad (10)$$

in which the constants α, β, γ, δ and ε are all positive. Since V_{min}, $V_{S,min}$ and $\overline{V}_S^-$ are negative, this means that each term except ε is promoting solvation. Equation (10) is one of the few GIPF relationships that include V_{min}, the overall most negative value of the electrostatic potential. The correlation coefficient for Eq. (10) is R = 0.988; the average absolute deviation from experiment is 0.27 kcal/mole for $\Delta G_{solvation}$ values varying over 9.59 kcal/mole.

The involvement in Eq. (10) of the site-specific $V_{S,max}$, $V_{S,min}$ and V_{min} indicates the importance of attractive interactions with, respectively, the oxygen and particularly the hydrogens of water, especially for the more polar solutes. The global $A_S^- \overline{V}_S^-$ terms are needed for the less polar ones; indeed, for a group of seven aromatic hydrocarbons, the free energy of solvation in water correlates well with the product $A_S^- \overline{V}_S^-$ alone. The role of its reciprocal in Eq. (10) appears to be to provide a minor correction. It is interesting that aqueous solvation depends so much upon the negative portions of the solute surfaces, showing that the attractive interactions with the hydrogens of water are much more significant than with the oxygen.

For three nonpolar solvents (cyclohexane, hexadecane and benzene), solvation is promoted by the solute having a large total surface area and large Π:[47]

$$\Delta G_{solvation} = -\alpha A_S{}^{0.5} - \beta A_S \Pi + \gamma A_S^+ \nabla_S^+ + \delta \qquad (11)$$

(B3P86/6-31+G**; α, β, γ, $\delta > 0$)

R varies between 0.933 and 0.945, with average absolute deviations from experiment of 0.23 - 0.26 kcal/mole for $\Delta G_{solvation}$ covering ranges of 3.1 kcal/mole. Thus electrostatic interaction is important, but not in a site-specific manner. Strongly positive regions on the solute, reflected in the product $A_S^+ \overline{V}_S^+$, have a hindering effect, presumably due to repulsion by the hydrogen peripheries of the solvent molecules. Finally, for solvents of intermediate polarity, solvation involves a combination of both site-specific and global features of the solute molecular surfaces,[47] depending on the types of interactions that dominate in each case. However large solute surface area consistently enhances solvation.

(*iv*) *Partition Coefficients*

Octanol/water partition coefficients, P_{ow}, which measure the relative solubilities of solutes in octanol and in water, are widely used as descriptors in quantitative structure-activity relationships (QSAR), for example in pharmacological and toxicological applications.[49] Since experimental values of these are not always available, a number of procedures for predicting them have been proposed (see references in Brinck et al.).[50]

We have shown, for 67 organic solutes of various chemical types, that log P_{ow} can be expressed by Eq. (12):[50]

$$\log P_{ow} = \alpha A_S - \beta \sigma_-^2 - \gamma A_S \Pi - \delta \qquad (12)$$

(HF/STO-5G*//HF/STO-3G*; α, β, γ, $\delta > 0$)

For Eq. (12), R = 0.961 and the average absolute deviation is 0.33 for log P_{ow} between −1.38 and 5.18. As was found in the case of solvation free energies (Section II.2.iii), large total area is a key to solubility in solvents of low polarity (in this case octanol). A high

σ_-^2 is the primary factor promoting the aqueous phase. This draws attention to the significant role of hydrogen bonding, with the water being the donor. These conclusions were reached earlier by Kamlet et al.[51] and by Famini et al.[52] A large Π (internal charge separation) also helps water solubility, but to a lesser extent than σ_-^2.

Two of the worst outliers were N,N-dimethylformamide and N,N-dimethyl-acetamide. For both of these, solubility in water was greatly underestimated. This may illustrate a situation in which conformation does assume importance. In the gas phase structures used to compute the surface properties, the nitrogens are planar. There is reason to believe, however, that interaction with water molecules will cause the nitrogens to become pyramidal,[48] since that produces more localized lone pairs that better attract water hydrogens. Thus, analysis involving planar nitrogens would not indicate the true strength of the interaction.

The relationship described for log P_{ow} was obtained for a diverse set of organic solutes. It can be expected to change in form when all are of a given chemical type; this was found for a group of nine nitroaromatics (plus the parent benzene and toluene).[53] Due to the presence of the electron-withdrawing nitro group, nitroaromatics are characterized by positive electrostatic potentials above and below the rings.[36] Thus, while surface area continues to favor octanol, as is shown in Eq. (13), it is now σ_+^2 and again Π that help water solubility.

$$\log P_{ow} = \alpha A_S - \beta A_S \Pi - \gamma \sigma_+^2 - \delta \qquad (13)$$

(HF/STO-5G*//HF/STO-3G*; α, β, γ, $\delta > 0$)

R = 0.993 and the standard deviation is 0.065; log P_{ow} varies over 1.55 units. The similarity between Eqs. (12) and (13) should be noted. They differ only in that the σ_-^2 term in the former, which reflects the importance of negative regions on the solutes interacting with water hydrogens, is replaced by σ_+^2, because it is

now the positive regions above and below the nitroaromatic rings that play a key role, interacting with the water oxygen.

(*v*) *Fullerene Solubility*

Given the interest in extended carbon systems in recent years, it seemed useful to study the solubility of C_{60} (fullerene) in various organic liquids.[54,55] It was now for the solvents that the molecular surface properties were computed. The resulting Eq. (14) shows that, for this large nonpolar solute, solubility is enhanced by solvent molecule surface area and by the latter having somewhat weak but relatively balanced (the product $\nu\sigma_{tot}^2$) positive and negative potentials, as does fullerene itself.

$$\log(\text{solubility}) = \alpha A_S^{\,4} + \beta(\nu\sigma_{tot}^2)^{0.5} - \gamma\frac{\sigma_{tot}^2}{A_S^{\,1.5}} - \delta \quad (14)$$

(HF/STO-5G*//HF/STO-3G*; $\alpha, \beta, \gamma, \delta > 0$)

R = 0.954, with an average absolute deviation of 0.338 over a range of 4.987.

(*vi*) *Diffusion Coefficients*

Diffusion coefficients provide a contrast to the preceding discussion, because now the focus is upon avoiding, rather than promoting, attractive solute-solvent interactions, which would hinder diffusion. For a series of photographic developing agents, all derivatives of benzene, the rate of diffusion through gelatin was found to decrease with surface area and σ_+^2 and to increase with σ_-^2.[56]

$$\text{Diffusion coefficient} = \alpha A_S^{\,-1} - \beta\sigma_+^2 + \gamma\sigma_-^2 - \delta \quad (15)$$

(HF/STO-5G*//HF/STO-3G*; $\alpha, \beta, \gamma, \delta > 0$)

R = 0.990, and the average absolute deviation is 0.065 in dry gelatin (range = 1.53) and 0.6 in wet (range = 13.8). The inverse dependence upon area, which promotes intermolecular attractions, is to be expected. The roles of σ_+^2 and σ_-^2 can be attributed to the medium, gelatin, which is a mixture of water-soluble proteins and thus contains a large number of >N–C(=O)– peptide linkages. The strongly-negative lone pair electrostatic potentials of the peptide nitrogen and oxygen will impede diffusion by attracting any positive sites on the moving molecules, but will repel negative sites. Diffusion is therefore favored by a high solute σ_-^2 but hindered by a high σ_+^2.

(*vii*) *Solubility in Supercritical Solvents*

Supercritical solutions are characterized by very low solvent densities. As a result, they possess the interesting feature that solubility is determined more by solute-solute than solute-solvent interactions. Thus we were able to express the solubilities of naphthalene and a series of indole derivatives in four different supercritical solvents (C_2H_4, C_2H_6, CO_2 and the highly polar CHF_3) in the same functional format, only the numerical coefficients varying from one to another.[57] Solute-solvent interactions do occur,[58] but solubility can be represented quite well by relating it inversely to solute surface area and σ_{tot}^2,[59] large values of which can be expected to favor solute-solute attractive interactions. The correlation coefficients were 0.940 and 0.921 for two different pressures. An interesting application of these findings was in identifying realistic simulants of highly toxic compounds, to be used in investigating the proposed disposal of the latter through supercritical oxidation.[60]

(*viii*) *Solvent Effects Upon Chemical Reactions*

Gholami and Talebi have recently demonstrated that the GIPF technique can be utilized in analyzing the effects of solvents upon chemical reactions.[61] They focused upon the Diels-Alder reaction between cyclopentadiene and menthyl acrylate in 15 different

solvents. The $V_S(\mathbf{r})$ quantities were computed for the solvent molecules, and expressions were developed for the rate constants and both the *endo*/*exo* and diastereofacial selectivities in the various solutions. The first two properties depend upon $V_{S,max}$, the product $\nu\sigma_{tot}^2$ and (inversely) molecular volume (which Gholami and Talebi used instead of area); the last upon $V_{S,max}$ and Π. The effects of the solvents were interpreted in terms of their hydrogen-bond-donating ability and solvophobicity.

3. Summary

The GIPF approach is effective in correlating and predicting, with satisfactory accuracy, various solution properties. It can also yield insight into the physical factors that are involved. It should be mentioned that in developing the relationships that have been presented, our primary purpose was to demonstrate the effectiveness of the procedure, and not necessarily to obtain the best possible correlation. Thus it may be that those discussed could be improved somewhat. This could also be achieved by treating different classes of compounds separately (e.g., hydrocarbons, alcohols, amines, etc.). We have usually tried to be as general as possible.

The GIPF technique does not address directly the interaction between a solute and a solvent, since one or the other is always included only implicitly. Thus the treatment is static rather than dynamic; for instance, no account is taken of how the solute and solvent may affect (e.g., polarize) each other. The next Section will survey procedures that do, in one way or another, quantitatively involve both partners in the interaction.

III. PROCEDURES THAT DIRECTLY ADDRESS THE SOLUTE-SOLVENT INTERACTION

1. Discrete Molecular Models of Solvent

(*i*) *Classical*

In principle, the ideal description of a solution would be a quantum mechanical treatment of the *supermolecule* consisting of representative numbers of molecules of solute and solvent. In practice this is not presently feasible, even if only a single solute molecule is included. In recent years, however, with the advances in processor technology that have occurred, it has become possible to carry out increasingly detailed molecular dynamics or Monte Carlo simulations of solutions, involving hundreds or perhaps even thousands of solvent molecules. In these, all solute-solvent and solvent-solvent interactions are taken into account, at some level of sophistication.

Molecular dynamics and Monte Carlo methodologies differ in that the former uses the equations of motion of classical physics to determine the movements of the atoms and arrive at the final state of the sytem after a lengthy series of time-steps, while the latter reaches it by random sampling of a large number of possible configurations, which are accepted or rejected on the basis of energetic criteria.[62-64] According to Orozco et al.,[11] molecular dynamics is preferable for large solutes, time-dependent processes and ease of setting up the simulation, while Monte Carlo is better for control of temperature and pressure, and reduction of configurational space.

A key to both methods is the force field that is used,[65] or more precisely, the inter- and possibly intramolecular potentials, from which can be obtained the forces acting upon the particles and the total energy of the system. An elementary level is to take only solute-solvent intermolecular interactions into account. These are typically viewed as being electrostatic and dispersion/exchange-repulsion (sometimes denoted van der Waals); they are represented by Coulombic and (frequently) Lennard-Jones expressions:

$$U_{\textbf{intermolecular}} = \frac{1}{4\pi\varepsilon_0}\sum_{\alpha}\sum_{\beta}\frac{Q_\alpha Q_\beta}{R_{\alpha\beta}} + \sum_{\alpha}\sum_{\beta}4\varepsilon_{\alpha\beta}\left[\left(\frac{\sigma_{\alpha\beta}}{R_{\alpha\beta}}\right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{R_{\alpha\beta}}\right)^{6}\right] \quad (16)$$

Sometimes the Buckingham potential is used, rather than the Lennard-Jones:

$$U_{\textbf{intermolecular}} = \frac{1}{4\pi\varepsilon_0}\sum_{\alpha}\sum_{\beta}\frac{Q_\alpha Q_\beta}{R_{\alpha\beta}} + \sum_{\alpha}\sum_{\beta}\left[A_{\alpha\beta}\exp(-B_{\alpha\beta}R_{\alpha\beta}) - C_{\alpha\beta}R_{\alpha\beta}^{-6}\right] \quad (17)$$

In Eqs. (16) and (17), α and β refer to atoms of the solute and solvent, respectively; ε_0 is the permittivity of free space, Q_α and Q_β are atomic charges, and $R_{\alpha\beta}$ is the distance between atoms α and β. The parameters $\varepsilon_{\alpha\beta}$, $\sigma_{\alpha\beta}$, $A_{\alpha\beta}$, $B_{\alpha\beta}$ and $C_{\alpha\beta}$ can either be assigned by fitting to experimental data or can be the arithmetic or geometric means of literature values for the individual atom types.[10,65,66] The atomic charges are commonly determined by requiring that they reproduce the calculated molecular electrostatic potentials.[10] In order to provide better descriptions of the solvent's structure, Eqs. (16) and (17) are generally extended to include solvent-solvent intermolecular interactions.

If it is desired to know how certain molecular properties of the solute (e.g., conformations) are affected by the presence of the solvent, then it is necessary to augment Eqs. (16) and (17) by appropriate solute intramolecular potentials. These would account for stretching, bending and torsional motions, plus any others deemed significant; for typical formulations, see Kollman,[10] Maple,[61] and Politzer and Boyd.[67] Equations (16) and (17) would also be expanded to encompass solute intramolecular interactions.

(*ii*) *Quantum-Mechanical/Classical*

The inter/intramolecular potentials that have been described may be viewed as classical in nature. An alternative is a hybrid quantum-mechanical/classical approach, in which the solute molecule is treated quantum-mechanically, but interactions involving the solvent are handled classically. Such methods are often labeled QM/MM, the MM reflecting the fact that classical force fields are utilized in molecular mechanics. An effective Hamiltonian H^{eff} is written for the entire solute/solvent system:

$$H^{eff} = H^{QM}_{solute} + H^{MM}_{solvent} + H_{solute-solvent} \tag{18}$$

In Eq. (18), H^{QM}_{solute} is a quantum-mechanical Hamiltonian for the gas phase solute molecule. $H^{MM}_{solvent}$ represents solvent inter- and intramolecular interactions, and is likely to consist of terms analogous to those in Eqs. (16) and (17). Finally, $H_{solute-solvent}$ consists of (a) Coulombic expressions for the electrostatic interactions between the electrons and nuclei of the solute molecule and the atomic charges Q_β within the solvent molecules, and (b) Lennard-Jones or Buckingham functions for the dispersion/exchange-repulsion interactions between solute and solvent atoms. (For more explicit depictions of H^{eff}, see Gao[9] and Orozco et al.[11]) Obtaining the total energy of the solute/solvent system requires solving,

$$E_{system} = \left\langle \varphi_{solute,solution} \middle| H^{QM}_{solute} + H_{solute-solvent} \middle| \varphi_{solute,solution} \right\rangle + H^{MM}_{solvent} \tag{19}$$

at each step of the simulation (i.e., each different configuration of the system), presumably by a self-consistent-field iteration. In view of the costliness of this in terms of computer resources, it is often preferred to use a semiempirical Hamiltonian for the solute

molecule, e.g., AM1[68] or PM3.[69] Among other possibilities is Warshel's valence bond formulation.[70]

(*iii*) *Free Energy Perturbation Procedure (FEP)*

Once the energy of the solute/solvent system has been determined, using some variant of a classical or a QM/MM force field, it can be used to find the free energy of solvation, which is the property of real interest. This is given, at constant temperature and pressure or at constant temperature and volume, respectively, by Eqs. (20) and (21):

$$\Delta G_{solvation} = \Delta H_{solvation} - T\Delta S_{solvation} \tag{20}$$

$$\Delta A_{solvation} = \Delta E_{solvation} - T\Delta S_{solvation} \tag{21}$$

$\Delta E_{solvation}$ could be computed by,

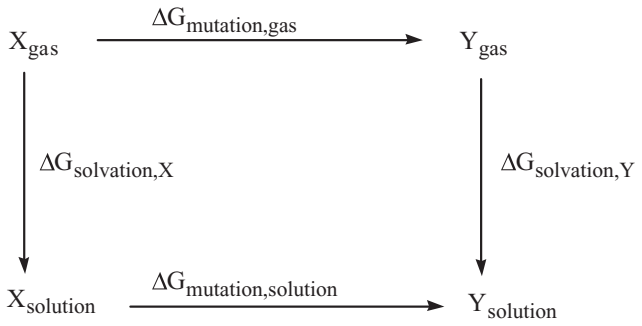
$$\Delta E_{solvation} = E_{system} - (E_{solute,gas} + E_{solvent}) \tag{22}$$

where,

$$E_{solute,gas} = \left\langle \varphi_{solute,gas} \middle| H^{QM}_{solute} \middle| \varphi_{solute,gas} \right\rangle \tag{23}$$

and $E_{solvent}$ is obtained by a simulation of the pure solvent. $\Delta H_{solvation}$ would then follow by converting $E_{solute,gas}$ into $H_{solute,gas}$ at the given temperature[71] and noting that for condensed phases, H ≈ E. Evaluating $\Delta S_{solvation}$ poses a problem, however, in applying Eqs. (20) and (21).[72]

A widely-used alternative is the free energy perturbation technique (FEP),[9-11,72-74] which will be discussed for the Gibbs free energy G. It involves a thermodynamic cycle such as the following, in which X and Y are two different molecules that are undergoing solvation in a particular solvent:



The horizontal portions of this cycle consist of X *mutating* reversibly into Y, in the gas phase and in solution. Since $\Delta G_{cycle} = 0$, then,

$$\Delta G_{solvation,Y} - \Delta G_{solvation,X} = \Delta G_{mutation,solution} - \Delta G_{mutation,gas} \qquad (24)$$

Thus if one can determine the free energy changes for the two mutations, then Eq. (24) will yield the free energy of solvation of Y relative to that of X, $\Delta\Delta G_{solvation,Y-X}$. This is accomplished through a sequence of simulations for which the force fields or Hamiltonians of X_{gas} and $X_{solution}$ are transformed by N small increments into those of Y_{gas} and $Y_{solution}$. $\Delta G_{mutation,gas}$ and $\Delta G_{mutation,solution}$ then result from applying a relationship from statistical mechanics, which in these instances takes the forms:

$$\Delta G_{mutation,gas} = -RT \sum_{i=0}^{N-1} \ln\left[\exp\left(-\frac{H_{i+1,gas} - H_{i,gas}}{RT} \right) \right] \qquad (25)$$

and

$$\Delta G_{mutation,solution} = -RT \sum_{i=0}^{N-1} \ln\left[\exp\left(-\frac{H_{i+1,solution} - H_{i,solution}}{RT} \right) \right] \qquad (26)$$

In Eqs. (25) and (26), the summations are over the incremental steps in going from X to Y in the gas phase or in solution. The H_i are the intermediate Hamiltonians (or force fields in a classical treatment). Thus, $H_{i=0,gas} = H_{X,gas}$, $H_{i=N,gas} = H_{Y,gas}$, etc. It is of course desirable that the molecules X and Y be structurally similar, so that the *perturbation* of X that produces Y be small. Another option is to let Y be composed of noninteracting (*dummy*) atoms,[75] so that its free energy of solvation is zero. Then Eq. (24) gives the absolute free energy of solvation of X:

$$\Delta G_{solvation,X} = \Delta G_{mutation,gas} - \Delta G_{mutation,solution} \qquad (27)$$

(*iv*) *Discussion of Discrete Molecular Models*

Table 1 lists some relative $\Delta G_{solvation}$ calculated with classical force fields. The average absolute deviation from experiment is 0.92 kcal/mole, the largest discrepancies involving some amines and acetamides;[75] some improvement was obtained for the amines when the force field included a term representing the polarization of the solute.[79] With a QM/MM technique, using the AM1 solute Hamiltonian,[68] the average absolute deviation for the systems in Table 2 is 1.5 kcal/mole. Gao has mentioned some problems that may be encountered with the AM1 solute Hamiltonian, but his overall assessment is positive.[9] In Table 3 can be seen the consequences of two different ways of estimating atomic charges. Those derived from the molecular electrostatic potentials are clearly more effective than the Mulliken[81] in this instance.

Once free energies of solvation are available, other solution properties can be determined, such as solute conformations, pK_a values, electrode potentials, reaction energetics, etc.[9,10,82] For example, Reynolds applied *ab initio* (HF and MP2) QM/MM approaches to computing the electrode potentials in water of a group of quinines;[83] the average absolute deviations for the most stable conformations were 0.024 (HF) and 0.033 (MP2) volts, for a range of 0.322 volts.

An important feature of QM/MM methods is that the polarization of the solute molecule's charge distribution by the solvent can be evaluated, since the wave function of the former is

Table 1
Relative Free Energies of Solvation in Water, in kcal/mole, Obtained by Classical Discrete Molecular Solvent Methods

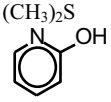
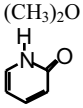
X	Y	$\Delta G_{solvation,Y} - \Delta G_{solvation,X}$		Ref.
		Calculated	Exp.	
C_2H_6	C_3H_8	0.2 ± 0.1	0.2	76
CH_3OH	C_2H_6	6.9 ± 0.1	6.9	76
CH_3OH	$(CH_3)_2O$	3.3	3.6	10
CH_3OH	CH_3NH_2	1.3 ± 0.1	0.5	77
CH_3OH	CH_3SH	3.5 ± 0.1	3.7	76
NH_3	CH_3NH_2	0.62 ± 0.05	–0.26	75
NH_3	$(CH_3)_3N$	4.36 ± 0.05	1.07	75
CH_3NH_2	$(CH_3)_2NH$	1.62 ± 0.01	0.27	75
$(CH_3)_2NH$	$(CH_3)_3N$	2.34 ± 0.02	1.06	75
$CH_3C(O)NH_2$	$CH_3C(O)NHCH_3$	2.09 ± 0.11	–0.40	75
$CH_3C(O)NHCH_3$	$CH_3C(O)N(CH_3)_2$	1.05 ± 0.02	1.53	75
$CH_3C(O)NHCH_3$	CH_4	11.6 ± 0.2	12.1	76
CH_3SH	C_2H_6	3.9 ± 0.1	3.1	77
CH_3SH	CH_3CN	–1.3 ± 0.2	–2.6	77
$(CH_3)_2S$	$(CH_3)_2O$	0.9 ± 0.1	0.4	76
		– 4.7 ± 0.4	– 4.1	78

Table 2
Relative Free Energies Of Solvation in Water, in kcal/mole,[a] Obtained by a QM/MM Discrete Molecular Solvent Method, Using The AM1 Solute Hamiltonian. The Estimated Error Bars for The Calculated Values Are ± 0.5 Kcal/Mole

X	Y	$\Delta G_{solvation,Y} - \Delta G_{solvation,X}$	
		Calculated	Experimental
C_2H_6	CH_4	1.2	0.2
C_2H_6	C_6H_6	0.3	–2.6
C_2H_6	H_2O	–8.3	–8.1
C_2H_6	CH_3OH	–6.2	–6.9
C_2H_6	$(CH_3)_2O$	–3.6	–3.7
C_2H_6	$(CH_3)_2CO$	–5.0	–5.6
C_2H_6	CH_3NH_2	–4.0	–6.4
C_2H_6	CH_3CN	–3.1	–5.7
C_2H_6	$CH_3C(O)NHCH_3$	–8.5	–12.0
C_2H_6	$CH_3CO_2^-$	–80	–79

[a] Reference 84.

obtained in both the gas and solution phases. The energy associated with this polarization is,[9]

$$E_{solute}^{polariz.} = \left\langle \varphi_{solute,solution} \middle| H_{solute}^{QM} + H_{solute-solvent} \middle| \varphi_{solute,solution} \right\rangle - \left\langle \varphi_{solute,gas} \middle| H_{solute}^{QM} \middle| \varphi_{solute,gas} \right\rangle - \left\langle \varphi_{solute,gas} \middle| H_{solute-solvent} \middle| \varphi_{solute,gas} \right\rangle \tag{28}$$

The last integral in Eq. (28) represents the interaction energy of the unperturbed gas phase solute with the solvent. The wave functions $\varphi_{solute,gas}$ and $\varphi_{solute,solution}$ also permit finding the dipole moments of the solute molecule in the gas and solution phases.

Table 3
Free Energies of Solvation in Water, in kcal/mole,[a] Obtained by Classical Monte Carlo Simulations Using Two Different Atomic Charge Definitions

Solute	$\Delta G_{solvation}$		$\Delta G_{solvation}$ (exp.)
	Mulliken charges	Electrostatic potential charges	
CH_3OH	–9.2 ± 0.5	–4.6 ± 0.4	–5.1
CH_3NH_2	–4.2 ± 0.5	–4.3 ± 0.5	–4.6
CH_3CN	–6.8 ± 0.5	–4.7 ± 0.5	–3.9
$(CH_3)_2O$	–11.3 ± 0.5	–1.4 ± 0.5	–1.9
CH_3SH	1.3 ± 0.5	0.0 ± 0.5	–1.2
CH_3Cl	0.5 ± 0.5	0.1 ± 0.5	–0.5
C_2H_6	3.0 ± 0.5	3.3 ± 0.5	1.8
$CH_3C(O)NH_2$	–13.5 ± 0.4	–13.4 ± 0.4	–9.7
$CH_3C(O)OH$	–10.3 ± 0.4	–8.5 ± 0.4	–6.7
$(CH_3)_2CO$	–5.7 ± 0.5	–3.6 ± 0.5	–3.8
$CH_3C(O)OCH_3$	–10.7 ± 0.5	–5.3 ± 0.5	–3.3
C_6H_6	–8.2 ± 0.4	–0.4 ± 0.4	–0.8
pyridine	–10.7 ± 0.4	–4.9 ± 0.4	–4.7
Ave. abs. deviation from exp.	**4.0**	**1.1**	

[a] Reference 80.

These results are given for the DNA bases in Table 4. The calculated effects are quite substantial, with the predicted increases in dipole moments essentially parallelling the magnitudes of the polarization energies. Given $\varphi_{solute,solution}$, molecular electrostatic potentials can also be obtained in solution, and used to derive atomic charges.[85,86]

The discrete molecular solvent methods that have been described can be quite effective for a variety of applications, and can provide insight into the polarization of the solute as well as the changes in solvent structure that occur in its vicinity. They do pose the problem of being computationally expensive, although this is increasingly mitigated by advances in software and in processor technology; Jorgensen pointed out already several years ago that computing the difference in the free energies of hydration of CH_3OH and C_2H_6, which required "two weeks on a high-end $100 000 minicomputer (Harris 80) in 1984, can be run today in a few hours on a $3000 Pentium II based PC."[74] Nevertheless, the continuum solvent methods to be discussed in the next Section still have the advantage of being less demanding in terms of computer resources.

Table 4
Dipole Moments (Debyes) in Gas Phase and in Aqueous Solution,[a] and Aqueous Polarization Energies (kcal/mole),[b] Computed by a QM/MM Discrete Molecular Solvent Method, Using The AM1 Solute Hamiltonian.

Molecule	$\mu_{gas,calc.}$	$\mu_{gas,exp.}$	$\mu_{aqueous}$	$E_{polarization}$
adenine	2.17	3.16	3.82	–3.8
cytosine	6.33	7.10	9.85	–8.5
guanine	6.18	6.76	8.47	–7.8
thymine	4.24	3.58	5.87	–4.0
uracil	4.28	3.86	5.85	–3.9

[a] Reference 84.
[b] Reference 9.

2. Continuum Models of Solvent

(*i*) *Classical Electrostatic Solvation Free Energy*

Continuum models are rooted in classical electrostatics, and its applications in the analysis of the dielectric constants of polar liquids. A key relationship is Poisson's equation,

$$\nabla^2 V(\mathbf{r}) = -\frac{4\pi}{\varepsilon} D(\mathbf{r}) \tag{29}$$

which links the electrostatic potential $V(\mathbf{r})$ to the total charge density $D(\mathbf{r})$ in a region with dielectric constant ε. $D(\mathbf{r})$ includes both electrons and nuclei. If $D(\mathbf{r}) = 0$ in the region, then Laplace's equation results,

$$\nabla^2 V(\mathbf{r}) = 0 \tag{30}$$

An early continuum treatment of solvation, associated with Born,[17] comes out of the analysis of the electrostatic work involved in building up a charge Q on a conducting sphere of radius R in a medium with dielectric constant ε. From Poisson's equation, it follows that the potential outside of the sphere is $Q/\varepsilon R$. Thus the work of charging is the result of each additional element dq interacting with the charge q already present:[87]

$$W_{\text{electrostatic}} = \int_0^Q \frac{q}{\varepsilon R} dq = \frac{1}{2}\frac{Q^2}{\varepsilon R} \tag{31}$$

The electrostatic component of the free energy of solvation of the sphere is then the difference between doing this charging in a vacuum ($\varepsilon = 1$) and in the medium:

$$G_{\text{electrostatic}} = -\frac{Q^2}{2R}\left(1 - \frac{1}{\varepsilon}\right) \tag{32}$$

Equation (32) has been quite useful for treating the solvation of monatomic ions (Section III.5).[88]

Since the effect of the charge Q outside of the sphere is as though it were located at the center, Eq. (32) can also be viewed as pertaining to a point charge (monopole) at the center of a spherical cavity. A few years later, it was extended to the case of a point dipole of magnitude μ,[18-20] which can of course be a model for neutral polar molecules:

$$G_{\text{electrostatic}} = -\frac{\varepsilon - 1}{2\varepsilon + 1}\left(\frac{\mu^2}{R^3}\right) \tag{33}$$

The continuum model of solvation has evolved from these beginnings. The solvent is treated as a continuous polarizable medium, usually assumed to be homogeneous and isotropic, with a uniform dielectric constant ε.[11-16] The solute molecule creates and occupies a cavity within this medium. The free energy of solvation is usually considered to be composed of three primary components:

$$\Delta G_{\text{solvation}} = G_{\text{electrostatic}} + G_{\text{cavitation}} + G_{\text{vdW}} \tag{34}$$

$G_{\text{electrostatic}}$ is the free energy of the electrostatic interaction between the solute molecule and the medium. This can be viewed in the following manner:[13] The charge distribution of the solute molecule induces polarization of the medium and the resulting induced field (the *reaction field*) in turn polarizes the molecule. This leads to a new reaction field, etc. The process continues until the work required to distort the solute and solvent charge distributions balances the free energy gained in the interaction. It can be anticipated that self-consistency will need to be established in the mathematical treatment. $G_{\text{cavitation}}$ is the free energy required to form the cavity, while G_{vdW} refers to the solute-medium van der Waals interaction; as discussed in Section III.1.i, G_{vdW} represents dispersion and exchange-repulsion effects. Equation (34) is often written with just $G_{\text{dispersion}}$ instead of G_{vdW}; it is also sometimes expanded to include a term that accounts for changes in solvent

structure in the neighborhood of the cavity.[13,16] However this is frequently taken to be part of $G_{cavitation}$.

The evaluation of $G_{electrostatic}$ has received a great deal of attention. It is clear that Eqs. (32) and (33), which are for nonpolarizable point charges and point dipoles, cannot reproduce the effect of the medium upon the solute molecule. A major contribution was made by Onsager, who took this molecule to be a polarizable point dipole located at the center of a spherical cavity;[20] the resulting expression is,

$$G_{electrostatic} = -\frac{\varepsilon - 1}{2\varepsilon + 1}\left(\frac{\mu^2}{R^3}\right)\left[1 - \frac{\varepsilon - 1}{2\varepsilon + 1}\left(\frac{2\alpha}{R^3}\right)\right]^{-1} \qquad (35)$$

R is the radius of the cavity, μ and α are the dipole moment and polarizability of the solute, and ε the dielectric constant of the solvent. Equation (35) does address the polarization of the solute molecule by the reaction field, although not carrying this to self-consistency. (It is interesting that Onsager's paper, the sixth-most-cited in the history of the *Journal of the American Chemical Society*, was rejected by the *Physikalische Zeitshrift*, to which it had initially been submitted.)[89]

While Onsager's formula has been widely used, there have also been numerous efforts to improve and generalize it. An obvious matter for concern is the cavity. The results are very sensitive to its size, since Eqs. (33) and (35) contain the radius raised to the third power. Within the spherical approximation, the radius can be obtained from the molar volume, as determined by some empirical means, for example from the density, the molar refraction, polarizability, gas viscosity, etc.[90] However the volumes obtained by such methods can differ considerably. The shape of the cavity is also an important issue. Ideally, it should be that of the molecule, and the latter should completely fill the cavity. Even if the second condition is not satisfied, as by a point dipole, at least the shape of the cavity should be more realistic; most molecules are not well represented by spheres. There was accordingly, already some time ago, considerable interest in progressing to more suitable cavities, such as spheroids[91,92] and ellipsoids,[93] using appropriate coordinate systems. Such shapes

still permit analytical solutions of Laplace's equation, which was invoked for point dipoles. With increasing computing power, and following Huron and Claverie's treatment of the mathematical problem,[94] more accurate irregular (*molecular*) cavities came into use, formed by overlapping spheres centered, for example, on the nuclei or chemical groups of the solute molecule, perhaps with van der Waals or some other appropriate radii. A seemingly reasonable approach would be to use an outer contour of the solute molecule's electronic density, as is done in the GIPF method (Section II.1). It has been found, however, that this is not effective in the present context.[16,95] A summary and a partial listing of some of the different cavity treatments have been given by Tomasi and Persico.[12]

Improvements have also been made in expressing the polarizing potential of the solute molecule. For a neutral molecule, the dipole term reflected in Eqs. (33) and (35) is just the first in the multipole expansion of the potential. The inclusion of more terms (which allows nonpolar molecules to be treated) was carried out already by Kirkwood,[19] and by others since that time, e.g., Rivail and Rinaldi and their colleagues,[96,97] whether through one-center or multi-center expansions.[12-14,98] A variant of this, the generalized Born model, invokes distributed monopoles,[99-101] centered on the nuclei of the solute molecule, and applies the Born formula, Eq. (32). The atomic charges can be computed quantum-mechanically, for example from the molecular electrostatic potential,[10] or they can be viewed as parameters and assigned by fitting to experimental data or from some standard source, such as Jorgensen and Tirado-Rives.[102]

The procedures that have been mentioned have focused upon Eqs. (32), (33) and (35), and have sought to depict more realistically the cavity and the polarizing effect of the solute molecule. The reverse polarization of the latter by the reaction field has often not been taken into account.[12] A technique that formulates $\Delta G_{electrostatic}$ in a different manner is the polarizable continuum model (PCM).[12,103,104] This represents the reaction field as coming from an *apparent surface charge*, a charge distribution $\sigma(\mathbf{r})$ spread over the cavity surface, which is determined by solving Poisson's equation. $\sigma(\mathbf{r})$ gives rise to a potential $V_\sigma(\mathbf{r})$,

$$V_\sigma(\mathbf{r}) = \int_S \frac{\sigma(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|} \tag{36}$$

where the integration is over the cavity surface. Then $G_{electrostatic}$ can be written either as (a) the work associated with building up this charge in the presence of the electrostatic potential due to the solute, $V_{solute}(\mathbf{r})$, a process mediated by the factor α,[87]

$$\begin{aligned} G_{electrostatic} &= \int_{\alpha=0}^{1} \int V_{solute}(\mathbf{r})\sigma(\mathbf{r})\alpha d\alpha d\mathbf{r} \\ &= \frac{1}{2}\int V_{solute}(\mathbf{r})\sigma(\mathbf{r})d\mathbf{r} \end{aligned} \tag{37}$$

or (b) the work of creating the total charge distribution of the solute molecule, $D_{solute}(\mathbf{r})$, in the presence of $V_\sigma(\mathbf{r})$:

$$\begin{aligned} G_{electrostatic} &= \int_{\alpha=0}^{1} \int V_\sigma(\mathbf{r})D_{solute}(\mathbf{r})\alpha d\alpha d\mathbf{r} \\ &= \frac{1}{2}\int V_\sigma(\mathbf{r})D_{solute}(\mathbf{r})d\mathbf{r} \end{aligned} \tag{38}$$

Since V_{solute} produces $\sigma(\mathbf{r})$,[12,103,104] and the resulting $V_\sigma(\mathbf{r})$ polarizes $D_{solute}(\mathbf{r})$, thereby giving rise to a new V_{solute}, it is clear that an iterative process could be implemented.[105,106]

(*ii*) *Quantum-Mechanical Electrostatic Solvation Free Energy*

The emphasis thus far has been upon classical approaches to evaluating $G_{electrostatic}$ (although quantum-mechanical methods might be used in some instances to obtain atomic charges or solute molecule electronic densities.) However continuum models lend themselves very well to quantum-mechanical treatments of solute-solvent interactions, and have indeed been incorporated into several standard codes, such as the Gaussian series.[29] $G_{electrostatic}$ is now given by,

$$G_{\text{electrostatic}} = \left\langle \varphi_{\text{solute,cavity}} \middle| H^{QM}_{\text{solute}} + V_R \middle| \varphi_{\text{solute,cavity}} \right\rangle - \left\langle \varphi_{\text{solute,gas}} \middle| H^{QM}_{\text{solute}} \middle| \varphi_{\text{solute,gas}} \right\rangle \quad (39)$$

in which H^{QM}_{solute} is again the Hamiltonian for the gas phase solute molecule and V_R is the perturbation operator corresponding to the solvent (i.e continuum) reaction field. The fact that V_R depends on $\varphi_{\text{solute,cavity}}$ imposes the requirement of self-consistency, as was already mentioned. The methods based on Eq. (39) are therefore labeled SCRF (self-consistent reaction field). Thus, quantum-mechanically, the polarization of the solute molecule can be fully described. Since only the solute molecule is considered explicitly in Eq. (39), it is assumed to be acceptable to neglect any error associated with not converting the two energies in Eq. (39) to free energies.

The form of V_R is generally related to one of the classical procedures that have been discussed.[11-14,107] There has been extensive use of truncated multipole expansions of both the solute molecule's charge distribution and the potential associated with the reaction field, the latter being expressed in terms of the former. Such procedures were pioneered by Rivail, Rinaldi et al.[14,108-110] at both semi-empirical and *ab initio* levels, initially for spherical and ellipsoidal cavities but later extended to irregular molecular shapes.[14,111,112] Both one-center and multi-center expansions have been utilized. An interesting comparison of these was given by Rivail and Rinaldi.[14] For formamide, at the Hartree-Fock 6-31G* level, one-, three- and six-center treatments all gave essentially the same $G_{\text{electrostatic}}$ (–9.13, –9.11 and –9.13 kcal/mole, respectively). For one center, the first-order (dipolar) term is the largest, –7.82 kcal/mole, but even the fifth-order is still significant, 0.48 kcal/mole. For three centers, third-order is sufficient to give –9.17 kcal/mole, while with six centers, second-order already produces –9.14 kcal/mole. The rate of convergence can also depend upon cavity shape; for *n*-propylamine, for instance, it was much more rapid in an ellipsoid than in a sphere.[96] For further discussion of convergence issues, see Tomasi and Persico[12] and Cramer and Truhlar.[13,16]

The generalized Born model (GBM) can be regarded as a special case of the preceding procedures; the reaction field is expressed in terms of a multi-center monopole representation of the solute molecule, using the Born formula, Eq. (32).[13,16,113-115] The centers are the atomic nuclei. The results are quite sensitive to the method used to calculate the atomic charges; Cramer and Truhlar, who have applied the GBM approach extensively[13,16,107] use their Class IV charges for this purpose.[16,107,116] Various techniques have been utilized to determine the radii.[16,95,101,107]

It is relatively straightforward to implement the polarizable continuum model (PCM) via Eq. (39).[11,12,105,117] The potential of the reaction field, $V_\sigma(\mathbf{r})$, is due to the ostensible (*virtual*) charge distribution $\sigma(\mathbf{r})$ on the cavity surface, which in turn is related to the potential $V_{solute}(\mathbf{r})$ that arises from the nuclei and electrons of the solute molecule, Eq. (2). Since the latter is likely to be further polarized by $V_\sigma(\mathbf{r})$, thus affecting $V_{solute}(\mathbf{r})$, iteration to self-consistency is needed,[105,106] as already has been pointed out. (However Montagnani and Tomasi suggest that this often has little practical consequence.)[118]

In Table 5 are compared values of $G_{electrostatic}$ predicted by these quantum-mechanical SCRF versions of the multipole expansion, GBM and PCM procedures.[107] The solvent was taken to have $\varepsilon = 80$ (for water, $\varepsilon = 80.20$ at 20 °C),[119] and molecular cavities were used. The computational level was HF/6-31G*. The multipole expansion $G_{solvation}$ tends to be smaller in magnitude than the GBM, and the PCM larger. The average absolute differences are 1.2 and 3.0 kcal/mole, respectively, although the latter would diminish to 2.2 kcal/mole if cytosine and guanine were not included. Curutchet et al. examine in detail the effects of using various atomic radii in establishing the cavities.[107] Barone et al. have also discussed this issue, and given a set of rules for determining atomic radii which, in a PCM treatment of hydration, decreased the mean errors in $\Delta G_{hydration}$ for 43 neutral molecules from 1.60 (Pauling radii)[119] and 0.91 (isodensity contour, $\rho(\mathbf{r}) = 0.004$ au) to 0.16 kcal/mole.[95] For molecular ions, the improvement was from 4.70 (Pauling radii) to 1.04 kcal/mole, for a range of −52 to −107 kcal/mole.

Table 5
Values of $G_{electrostatic}$ Predicted by Three Quantum-Mechanical SCRF Methods, in kcal/mole.[a] The Solvent Has ε = 80.

Solute	Method		
	Multipole expansion	GBM	PCM
CH_3F	–1.9	–3.5	–3.2
$CHCl_3$	–2.1	–1.1	–1.9
C_6H_5SH	–2.9	–2.4	–4.3
$CH_3C(O)OCH_3$	–3.8	–5.5	–6.8
$CH_3CH_2NO_2$	–5.7	–5.6	–8.3
$(CH_3)_2CO$	–4.2	–5.7	–7.5
CH_3CN	–5.5	–8.1	–7.4
C_6H_5CHO	–4.7	–5.8	–8.0
CH_3NH_2	–3.1	–1.7	–6.1
CH_3OH	–3.5	–3.8	–6.8
C_6H_5OH	–4.7	–4.5	–7.0
$CH_3C(O)OH$	–4.8	–6.5	–9.5
$CH_3C(O)NH_2$	–7.1	–8.6	–12.8
cytosine	–14.8	–16.6	–23.6
guanine	–16.7	–17.6	–27.5

[a] Reference 107.

(*iii*) *Cavitation and van der Waals Solvation Free Energies*

$G_{cavitation}$ and G_{vdW} are sometimes labeled as nonpolar, first-solvation-shell or short-range contributions to $\Delta G_{solvation}$. They are, in general, of the same order of magnitude as $G_{electrostatic}$, but since they have opposite signs ($G_{cavitation} > 0$, $G_{vdW} < 0$), they partially cancel, so that for polar solutes, normally $\left|G_{electrostatic}\right| > \left|G_{cavitation} + G_{vdW}\right|$. This will be seen in data to be presented in the next Section.

Various approaches to evaluating $G_{cavitation}$ and G_{vdW} have been proposed; for overviews, see Tomasi and Persico[12] and Cramer and Truhlar.[16] It seems reasonable to relate $G_{cavitation}$ to the area of the cavity and the surface tension of the solvent, as was suggested by Uhlig.[120] An observation made by Hermann indicated that this could be extended to include G_{vdW} as well; he showed that the free energies of hydration of a group of saturated hydrocarbons correlate linearly with their solvent-accessible surface areas

(SA).[121] If one assumes that $G_{electrostatic}$ is negligible for hydrocarbons, then Hermann's finding could be interpreted as implying that, approximately,

$$G_{cavitation} + G_{vdW} = \alpha(SA) + \beta \tag{40}$$

While Eq. (41) has had some success.[101,122] what is now widely used is a version that is specific to each individual atom:[107,123,124]

$$G_{cavitation} + G_{vdW} = \sum_i \alpha_i (SA)_i + \beta \tag{41}$$

In Eq. (41), $(SA)_i$ is the solvent accessible surface area of atom i. The parameters α_i and β are assigned by fitting to experimental $\Delta G_{solvation}$; this has the added benefit, as shall be seen, of helping to compensate for any systematic errors in $G_{electrostatic}$. Some researchers prefer to treat $G_{cavitation}$ and G_{vdW} separately,[107,125-127] perhaps utilizing a relation-ship analogous to Eq. (41) for either $G_{cavitation}$ or G_{vdW}, in combination with variants such as the Claverie-Pierotti expression for $G_{cavitation}$[128] or the Floris-Tomasi for G_{vdW}.[129,130]

(*iv*) *Discussion of Continuum Models*

Table 6 lists computed $G_{cavitation} + G_{vdW}$ and also $\Delta G_{solvation}$ for a group of solutes in a solvent with $\varepsilon = 80$, using molecular cavities. This table complements Table 5, which gave the corresponding $G_{electrostatic}$. The multipole expansion and GBM $G_{cavitation} + G_{vdW}$ were obtained from Eq. (41) (but with different parameters) while the PCM work used Eq. (41) for G_{vdW} but the Claverie-Pierotti $G_{cavitation}$.[128] The $\Delta G_{solvation}$ can be compared to the experimental free energies of hydration of these solutes, since water has $\varepsilon = 80.20$ at 20 °C.[119]

Several notable points are brought out by Table 6:

1. In view of the approximations associated with the calculations, the results—by all three procedures—are

Table 6
Values of $G_{cavitation} + G_{vdW}$ and $\Delta G_{solvation}$ Predicted by Three Quantum-Mechanical SCRF Methods, in kcal/mole.[a] The Solvent Has ε = 80.

Solute	Multiple expansion		GBM		PCM		ΔG_{hydr} (exp.)
	$G_{cav} + G_{vdW}$	ΔG_{solv}	$G_{cav} + G_{vdW}$	ΔG_{solv}	$G_{cav} + G_{vdW}$	ΔG_{solv}	
CH_2F	—	—	2.5	–1.0	1.9	–1.3	–0.2
$CHCl_3$	—	—	0.4	–0.7	1.4	–0.5	–1.1
C_6H_5SH	—	—	–0.1	–2.5	2.8	–1.5	–2.6
$CH_3C(O)OCH_3$	0.1	–3.7	1.9	–3.6	4.1	–2.7	–3.3
$CH_3CH_2NO_2$	—	—	1.4	–4.2	4.8	–3.5	–3.7
$(CH_3)_2CO$	–0.2	–4.4	1.5	–4.2	3.4	–4.1	–3.9
CH_3CN	1.5	–4.0	2.6	–5.5	3.9	–3.5	–3.9
C_6H_5CHO	—	—	1.0	–4.8	3.2	–4.8	–4.0
CH_3NH_2	–1.9	–5.0	–2.7	–4.4	1.4	–4.7	–4.6
CH_3OH	–1.4	–4.9	–1.5	–5.3	1.1	–5.7	–5.1
C_6H_5OH	—	—	–2.3	–6.8	1.6	–5.4	-6.6
$CH_3C(O)OH$	–2.2	–7.0	–0.3	–6.8	2.6	–6.9	–6.7
$CH_3C(O)NH_2$	–3.0	–10.1	–2.0	–10.6	3.0	–9.8	–9.7
Cytosine	—	—	–4.1	–20.7	2.7	–20.9	—
Guanine	—	—	–5.9	–23.5	1.9	–25.6	—
Ave. abs. deviation from exp:		0.3		0.5		0.6	

quite good, with average absolute deviations from experiment between 0.3 and 0.6 kcal/mole.

2. The quantities $G_{cavitation} + G_{vdW}$, because of being parametrized by fitting to experimental $\Delta G_{solvation}$, do help to overcome discrepancies in $G_{electrostatic}$. As was mentioned in Section III.2.ii, the $G_{electrostatic}$ given by the multipole expansion and the GBM techniques differed by an average 1.2 kcal/mole, the PCM and the GBM by 3.0 kcal/mole. The corresponding values for the $\Delta G_{solvation}$ in Table 6 are 0.5 and 0.7 kcal/mole. Particularly striking are the similarities between the PCM and GBM $\Delta G_{solvation}$ for cytosine and guanine (Table 6), in contrast to the significant disagreements in their $G_{electrostatic}$ (Table 5).
3. The data in Tables 5 and 6 bear out the earlier statement that, for polar solutes, $G_{electrostatic}$ is normally larger in magnitude than $G_{cavitation} + G_{vdW}$, sometimes considerably so. The only exception to this in these two tables is CH_3NH_2, at the GBM level.

In a very extensive test of the SM5 method (a type of GBM), Hawkins et al. found the average absolute deviation in $\Delta G_{solvation}$ to be 0.38 kcal/mole for 260 molecular solutes in water and in 90 organic solvents;[131] for ions in water, it was 3.8 kcal/mole, for experimental $\Delta G_{solvation}$ between −58 and −110 kcal/mole. Semiempirical quantum-mechanical procedures were used.

Table 7 compares free energies of hydration[125] produced by the two types of solvent models that have been presented: discrete molecular and continuum. The discrete molecular involved classical force field molecular dynamics (MD) and a free energy perturbation (FEP) technique whereby the solute molecule is *annihilated* to dummy atoms, so that absolute $\Delta G_{hydration}$ are obtained; the continuum were SCRF/PCM calculations, with Claverie-Pierotti $G_{cavitation}$ and Floris-Tomasi G_{vdW}. The computational level was the Hartree-Fock 6-31G*. The SCRF/PCM results are somewhat better than the MD/FEP, although the average absolute deviation of the latter would decrease to 0.9 kcal/mole if $CH_3C(O)OH$ were excluded. Orozco et al. analyze in detail the possible sources of error in the two computational approaches.[125]

Table 7
Free Energies of Hydration Predicted by Discrete Molecular (MD/FEP) and Continuum (SCRF/PCM) Solvent Models, in kcal/mole.[a]

Solute	Method		
	MD/FEP	SCRF/PCM	Experimental
NH_3	–5.6	–4.7	–4.3
H_2O	–5.8	–6.3	–6.3
HCN	–2.8	–3.9	–3.2
CH_3OH	–3.6	–4.1	–5.1
CH_3NH_2	–4.4	–3.3	–4.6
$CH_3C(O)OH$	–10.0	–5.8	–6.7
CH_3CHO	–4.6	–4.5	–3.5
$(CH_3)_2CO$	–4.9	–4.6	–3.9
Ave. abs. deviation from exp.	1.2	0.8	

[a] Reference 125.

There have been numerous applications of continuum models to equilibria and reactions in solution; surveys of these and extensive listings are provided by Cramer and Truhlar.[16] Other studies have focused upon the effects of solvents upon solute molecular properties, such as electronic and vibrational spectra,[16] dipole moments, nuclear quadrupole and spin-spin coupling constants and circular dichroism.[12]

Baik and Friesner used an SCRF procedure in conjunction with B3LYP density functional calculations to obtain electrode potentials for groups of aromatic molecules, metallocenes and transition metal complexes in four different solvents;[132] the average absolute deviation was about 0.15 volts for a range of values of 3.82 volts. Saracino et al. computed pK_a for a series of carboxylic acids with an average absolute deviation of 0.41 for pK_a between 1.23 and 5.03.[133]

It is evident that continuum models can be quite effective, for ionic solutes as well as for neutral ones. They also have the advantage of not being highly demanding in terms of computer resources. However a problem associated with these methods is posed by so-called *first-solvation-shell effects*.[16] One aspect of this is the difficulty of properly accounting for specific types of solute-

solvent interactions that may be present, such as hydrogen bonding. Another relates to the fact that the properties of the solvent, such as its dielectric constant and structure, are not the same in the immediate vicinity of the solute molecule as in the bulk. To some extent, such effects can be addressed through the empirical parametrization of $G_{cavitation}$ and G_{vdW}, and through the techniques used to assign solute atomic charges and radii; these are important issues.[95,107,134,135] An alternative is to use a hybrid or intermediate procedure, that combines features of both discrete molecular and continuum solvent models. Some examples will be discussed in the next Section.

3. Hybrid/Intermediate Models of Solvent

One means of trying to minimize the problems mentioned at the end of the last Section is to describe the first solvation shell (or perhaps the first two) as composed of discrete molecules, with which the interactions are treated explicitly (for instance by molecular dynamics or Monte Carlo simulations, or by cluster calculations) and then to use a continuum model for the remainder of the solvent.[16] This makes it possible to deal more adequately with hydrogen bonding[136,137] and other inner-solvation-shell factors. Good results have been obtained by hybrid techniques, for ions as well as for neutral solutes.[138-142] A very recent example is the approach proposed by Leontyev et al.,[143,144] which was designed for ions and strongly-polar molecules in polar solvents. The solute cavity is surrounded by a shell containing explicit point charge solvent particles; beyond this is the continuum solvent. The cavity, shell and continuum each have a different dielectric constant. The interaction of the solute with the shell and the continuum are treated separately, the first via molecular dynamics simulation. The results for a set of polyatomic ions will be discussed in Section III.5.iv.

A different kind of hybrid procedure was utilized by Goncalves and Stassen[145] to calculate $\Delta G_{solvation}$ for a series of solutes in water and three organic solvents. A continuum-based (PCM) $G_{electrostatic}$ was combined with the Claverie-Pierotti $G_{cavitation}$ and G_{vdW} produced by molecular dynamics simulation. The resulting $\Delta G_{solvation}$ are quite good, as can be seen from the

sample presented in Table 8, which also lists the contributions of the three components. In these solvents of low dielectric constant (at 20°C, $\varepsilon = 4.806$ for $CHCl_3$ and 2.284 for benzene),[119] G_{vdW} is usually larger in magnitude than $G_{electrostatic}$.

A solvent model that is intermediate between discrete molecules and continuum is that based upon *Langevin dipoles*.[146-148] The solute is surrounded by a cubic grid of polarizable point dipoles, which represent the solvent. Each feels the electric field due to the solute charges and the other dipoles, and its orientation and new dipole moment in response to this field are determined iteratively using a Langevin-type formula. (The Langevin equation was developed to treat Brownian movement.)[149] A noniterative technique, using a position-dependent screening function, is also possible. The solute then interacts with the potential due to these dipoles. Good results have been obtained for $\Delta G_{solvation}$.[148]

Another procedure that can be viewed as intermediate in nature is the finite differences method (FDM).[15,150-153] A three-dimensional grid is introduced into the entire system, including the solute cavity. The dielectric constant may be uniform (although different for solute and solvent) or position-dependent. A charge is assigned to each grid point; since the latter may not coincide with the positions of the solute nuclei, its atomic charges must be distributed to the neighboring grid points. A Poisson equation is solved numerically and iteratively at each grid point, using finite differences, to give the reaction field potential with which the solute interacts. The accuracy of the resulting $\Delta G_{solvation}$ depends in part upon the number of grid points; the spacing should be 0.25 A–0.50 A. In the finite elements method (FEM),[15,154] a grid is established in the solvent but the focus now is on the volume elements that it creates; to each of these corresponds a point charge equal in magnitude to the charge density in the element times its volume.

Table 8
Values of $\Delta G_{solvation}$ and Its Components in Two Different Solvents, in kcal/mol, As Predicted by a Hybrid Discrete Molecular/Continuum Solvent Model[a]

Solute	G_{cav}	G_{vdW}	$G_{electrostatic}$	$\Delta G_{solvation}$	$\Delta G_{solvation}$ (exp)
		Solvent: $CHCl_3$			
n-octane	18.55±0.71	–16.88±0.66	–7.54±0.41	–5.88±0.73	–5.25
benzene	12.05±0.43	–11.77±0.93	–5.41±0.50	–5.13±0.20	–4.64
n-propanol	10.94±0.73	–8.34±0.43	–6.70±0.76	–4.11±0.23	–4.41
1,4-dioxane	11.99±0.47	–11.36±0.85	–6.54±0.35	–5.91±0.18	–6.21
pyridine	11.88±0.80	–9.05±0.92	–8.79±0.47	–5.97±0.57	–6.45
ammonia	5.99±0.45	–4.27±0.55	–3.88±0.66	–2.17±0.21	–2.41
water	5.37±0.35	–4.06±0.44	–4.09±0.63	–2.79±0.49	–2.05
		Solvent: benzene			
n-octane	18.67±0.20	–18.92±0.44	–6.75±0.66	–6.95±0.73	–5.35
benzene	11.61±0.87	–7.77±0.83	–8.84±0.65	–5.00±0.64	–4.55
n-propanol	10.65±0.66	–9.42±0.85	–5.51±0.83	–4.28±0.99	–3.87
1,4-dioxane	11.66±0.19	–10.33±0.76	–7.15±0.46	–5.81±0.75	–5.21
pyridine	12.31±0.55	–11.50±0.44	–7.09±0.84	–6.31±0.88	–5.28
ammonia	4.78±0.17	–3.88±0.75	–2.74±0.73	–1.84±0.41	–1.12
water	4.64±0.99	–3.24±0.46	–3.41±0.25	–2.02±0.44	–1.71

4. General Comments: Discrete Molecular and Continuum Solvent Models

Sections III.1-III.3 have described some basic discrete molecular and continuum treatments of solute-solvent interactions. There are many variants and refinements of these that have not been discussed, such as the use of effective dielectric constants[66] or the implementation of dielectric screening.[155,156] For ionic solutions, it is sometimes preferred to find the reaction field potential via the Poisson-Boltzmann rather than the Poisson equation,[132,157] since the effects of the other ions can readily be incorporated into the former.[158]

The level of accuracy that can be achieved by these different methods may be viewed as somewhat remarkable, given the approximations that are involved. For relatively small organic molecules, for instance, the calculated $\Delta G_{solvation}$ is now usually within less than 1 kcal/mole of the experimental value, often considerably less. Appropriate parametrization is of key importance. Applications to biological systems pose greater problems, due to the size and complexity of the molecules,[66,156,159-161] and require the use of semiempirical rather than *ab initio* quantum-mechanical methods. In terms of computational expense, continuum models have the advantage over discrete molecular ones, but the latter are better able to describe solvent structure and handle first-solvation-shell effects.

5. Ionic Solvation

(*i*) *Experimental Data*

Before considering different theoretical approaches to determining the free energies and other thermodynamic properties of ionic solvation, it is important to be aware of a problem on the experimental level. There are several methods available for obtaining these quantities for electrolyte solutions, both aqueous and nonaqueous; some of these have been described by Conway and Bockris[162] and by Padova.[163] For example, enthalpies of solvation can be found via thermodynamic cycles, free energies from solubilities or galvanic cell potentials. However the results

pertain to the electrolyte as a whole, and there is no rigorous basis for apportioning them between the positive and negative ions. (It is the same problem that is encountered with respect to individual ionic activity coefficients.)

Various schemes have been devised for obtaining the thermodynamic solvation properties of the separate cations and anions from the measured electrolyte values.[162,163] Some have been based upon the concept that the solvation of small ions is primarily electrostatic in nature and its strength is inversely proportional to the ionic radius. Bernal and Fowler suggested that for a salt having positive and negative ions with similar crystal radii, the enthalpy of hydration could be divided essentially equally between them;[164] then these could be found for other ions from linear relationships between ionic $\Delta H_{hydration}$ and the estimated ion-water interaction energies. Latimer et al.[165] and Verwey[166] proposed techniques that involved plotting the differences in, respectively, the free energies and enthalpies of hydration for series of alkali halides having a common cation or anion against the inverse ionic radii. For a critique of these approaches, see Conway and Bockris.[162] Extension to other solvents was sometimes carried out by assuming that the ratio of the values for a given ion pair remains constant.[167]

More recently, ionic enthalpies and free energies of hydration have been determined relative to H^+.[88,168-171] This can be done, for example, by means of appropriate thermodynamic cycles.[170,171] Then by assigning the $\Delta H_{hydration}$ and $\Delta G_{hydration}$ of H^+ on some basis, these quantities can be evaluated for the other ion. Rashin and Honig took $\Delta H_{hydration}$ (H^+) to be −262.18 kcal/mole,[88] while Florian and Warshel assumed that $\Delta G_{hydration}$ (H^+) is −259.5 kcal/mole.[171]

Table 9 compares ionic enthalpies of hydration from the Bernal and Fowler,[164] Latimer et al.[165] and Rashin and Honig[88] procedures. Given the inherent uncertainty, the latter two sets of data are remarkably similar, considering that they were obtained 46 years apart. A number of tabulations of the thermodynamic solvation properties of ions in various solvents have now appeared.[88,163,168-171] It is important to keep in mind, however, that there is a degree of arbitrariness associated with the *experimental* $\Delta H_{solvation}$ and $\Delta G_{solvation}$ of individual ions.

Table 9
Experimental Ionic Enthalpies of Hydration Obtained By Three Different Methods, in kcal/mole

Ion	$\Delta H_{hydration}$		
	Bernal and Fowler[a]	Latimer et al.[b]	Rashin and Honig[c]
Li^+	−136	−121.2	−125.8
Na^+	−114	−94.6	−98.5
K^+	−94	−75.8	−78.4
Rb^+	−87	−69.2	−73.4
Cs^+	−80	−62.0	−67.6
F^-	−97	−122.6	−119.3
Cl^-	−65	−88.7	−85.3
Br^-	−57	−81.4	−78.8
I^-	−47	−72.1	−69.0

[a] Reference 164
[b] Reference 165
[c] Reference 88

(*ii*) *Ionic Solvation via Discrete Molecular Solvent Models*

Techniques such as those described in Section III.1, involving molecular dynamics or Monte Carlo simulations, have been used extensively to study ionic solvation.[84,157,172-176] A classical intermolecular potential, e.g., Eq. (16), is often used for both ion-solvent and solvent-solvent interactions. The number of solvent molecules utilized to create the ion's environment can have a significant effect upon the computed energetics;[172-174] an extreme case is that of F^- in water, for which $\Delta H_{hydration}$ changed from −111 to −146 kcal/mole in going from 125 to 216 water molecules.[173] This was attributed to a lesser degree of solvent disruption in the larger system. In general, 216 solvent molecules seems to be reasonable for monovalent ions, although more may be needed for divalent ones.

Solvent reorganization can indeed be an important component of ionic solvation. For their aqueous solutions, Chandrasekhar et al. found that this required an energy input that was, in magnitude, from 25% to 50% of the ion-solvent attraction.[173]

Since the solute-solvent interactions are now stronger (especially for polar solvents) than when the solute is molecular,

the capability of the simulation methodology, and the discrete molecular solvent model, for revealing solvent structure can be particularly useful and important. Thus, for a variety of ions in water (Li^+, Na^+, K^+, Rb^+, Ca^{++}, Sr^{++}, F^-, Cl^-, OH^-), there have been reported the coordination numbers, extents of hydrogen bonding (both ion-solvent and solvent-solvent), radial distributions of solvent molecules around the ion, etc.[172,173,175,176] It has further been possible to estimate the separate contributions of the first and outer solvent shells to the total free energy of solvation.

In Table 10 are listed some computed ionic enthalpies and free energies of hydration, as well as the corresponding experimental ones. The range in some of the latter illustrates the uncertainty associated with the experimental data, discussed in the previous Section. This must of course be taken into account in evaluating the calculated results.

The large discrepancies in $\Delta H_{hydration}$ for Li^+ and Na^+ have been blamed in part upon failure to adequately describe the polarization of the first solvation shell by the strong electric fields of these small ions.[173,175] Another source of error is the finite size of the simulation cell, which makes it necessary to cut off the interactions at a certain distance from the ion, e.g., 8.5 A. Corrections for this problem have been proposed.[174,175,179]

(*iii*) *Ionic Solvation via Continuum Solvent Models*

A straightforward approach to ionic solvation, particularly for monatomic ions, is by means of the Born equation,[17] Eq. (32), which was introduced in Section III.2.i:

$$G_{electrostatic} = -\frac{Q^2}{2R}\left(1-\frac{1}{\varepsilon}\right) \qquad (32)$$

This gives the electrostatic contribution to the free energy of solvation of a conducting sphere with charge Q and radius R in a medium with dielectric constant ε.

Equation (32) can be viewed as a forerunner of the cavity/continuum models of solvation that were discussed earlier

(Section III.2). As was pointed out, the total free energy of solvation is obtained by combining $G_{electrostatic}$ with the free energy needed to create the cavity plus that reflecting the solute-solvent van der Waals interaction:

$$\Delta G_{solvation} = G_{electrostatic} + G_{cavitation} + G_{vdW} \quad (34)$$

The shape and dimensions of the cavity are a key issue. In the case of Eq. (32), it is spherical, but what value should be assigned to R? For a monatomic ion, an obvious choice is its

Table 10
Calculated (By Simulations) And Experimental Ionic Enthalpies And Free Energies of Hydration, in kcal/mole

Ion	$\Delta H_{hydration}$		$\Delta G_{hydration}$	
	Calculated[a]	Experimental	Calculated[b]	Experimental
Li^+	−165 ± 4	−123.0[c]		
		−125.8[d]		
		−130[e]		
Na^+	−126 ± 4	−96.9[c]	−76.7	−89.6[g]
		−98.5[d]		−100[f]
		−102[e]		
		−106[f]		
F^-	−111 ± 4	−116[e]		
		−119.3[d]		
Cl^-	−80 ± 5	−81.4[f]	−70.7	−77[h]
		−82[e]		−77.7[f]
		−85.3[d]		−78[i]
OH^-			−93.3	−106[h]
				−110[i]

[a] Reference 173
[b] Reference 176. The simulations were at constant volume; accordingly these are Helmholtz free energies, $\Delta A_{hydration}$. For these aqueous solutions at 298 K, however, it is anticipated that $\Delta A_{hydration} \approx \Delta G_{hydration}$
[c] Reference 169
[d] Reference 88
[e] Reference 177
[f] Reference 168
[g] Reference 178
[h] Reference 170. The original values have been converted to a 1 molal ideal solution standard state (see reference 176)
[i] Reference 171

crystallographically-determined ionic radius; however this was found to exaggerate the magnitude of $\Delta G_{solvation}$.[165] (This may have been because $G_{continuum}$ + G_{vdW} were neglected; while these partially cancel, their sum is often opposite in sign to $G_{electrostatic}$, as can be seen in Tables 6 and 8.)

In a detailed study, Rashin and Honig concluded that crystal radii are suitable for anions but that cations should be given the covalent radii of the corresponding neutral atoms.[88] However they still found it necessary to increase both by an empirical 7% in order to satisfactorily reproduce experimental data. Rashin and Honig used the Gibbs-Helmholtz equation[90] to convert Eq. (32) into,

$$\Delta H_{electrostatic} = -\frac{Q^2}{2R}\left[1-\frac{1}{\varepsilon}-\frac{T}{\varepsilon^2}\left(\frac{\partial\varepsilon}{\partial T}\right)\right] \qquad (42)$$

and taking $(T/\varepsilon)(\partial\varepsilon/\partial T) = -1.357$ for water at 298 K,[180] they calculated $\Delta H_{electrostatic}$ for 28 monatomic ions. Some of their results are given in Table 11, along with *absolute* experimental values obtained by assigning $\Delta H_{hydration}$ for H^+ to be −262.18 kcal/mole.[88] For all 28 ions, the calculated $\Delta H_{electrostatic}$ differ from the absolute experimental $\Delta H_{hydration}$ by, on the average, 2.5%. This overall level of agreement indicates that Eq. (42), with appropriate R, is generally effective in predicting $\Delta H_{hydration}$ for monatomic ions. Rashin and Honig do warn that R for a given ion will vary from one solvent to another, but they believe that the values for water should be good initial approximations for other hydrogen-bonding solvents.

It was further pointed out by Rashin and Honig that the form of Eqs. (32) and (42) precludes significant dependence upon ε once it is greater than roughly 30.[88] However these equations also emphasize the importance of long-range interactions ($1/2R \approx 16$ kcal/mole for R = 10 A), and the authors correctly anticipate the problems encountered in simulations due to the limited sizes of the systems.[174,175, 179]

The Born equation does not take into account the mutually polarizing effect that the solute and solvent have upon each other. This can be done, within the framework of continuum solvent

models, by means of self-consistent reaction fields, at both classical and quantum mechanical levels. Such methods, for example the polarizable continuum and the generalized Born models (PCM and GBM) were described in Section III.2. Variations of these techniques have been applied to numerous monatomic and polyatomic ions.[95, 101, 131, 157]

Table 11
Ionic $\Delta H_{electrostatic}$ Obtained with The Born Equation In The Form of Eq. (42) Compared To Experimental $\Delta H_{hydration}$, In kcal/mole[a]

Ion	$\Delta H_{electrostatic}$, Eq. (42)	$\Delta H_{hydration}$, experimental
Li^{+}	–126.7	–125.8
Na^{+}	–99.3	–98.5
K^{+}	–76.8	–78.4
Rb^{+}	–72.2	–73.4
Cs^{+}	–66.3	–67.6
Mg^{+2}	–458.4	–462.4
Ca^{+2}	–358.3	–383.6
Sr^{+2}	–324.7	–348.3
Ba^{+2}	–314.8	–314.3
Al^{+3}	–1122.7	–1118.1
Ga^{+3}	–1122.1	–1124.1
In^{+3}	–935.1	–986.4
Sc^{+3}	–974.1	–939.8
Cu^{+}	–133.2	–143.5
Cu^{+2}	–532.8	–504.5
Zn^{+2}	–498.7	–491.6
Ag^{+}	–116.3	–115.1
Cd^{+2}	–442.1	–434.6
Hg^{+2}	–432.9	–439.4
F^{-}	–117.2	–119.3
Cl^{-}	–86.1	–85.3
Br^{-}	–79.9	–78.8
I^{-}	–71.2	–69.0
S^{-2}	–338.8	–325.0

[a] Reference 88

Assigning atomic, and hence cavity, sizes is always a key problem. Barone et al. developed a set of rules for determining individual atomic radii, that reflect their hybridization and their neighbors, with corrections to be applied when the atoms are in ions.[95] When these radii were implemented in a PCM procedure, the resulting $\Delta G_{hydration}$ for 27 ions differed from the experimental by an absolute average of only 1 kcal/mole. When Pauling radii[119] were used, this increased to 4.6 kcal/mole. Several of these PCM $\Delta G_{hydration}$ are presented in Table 12.

Table 12
Comparison of $\Delta G_{hydration}$, In Kcal/Mole, Produced By Specific Versions of PCM And GBM Methodologies.

Ion	$\Delta G_{hydration}$			
	PCM, Barone et al. radii[a]	PCM, Pauling radii[a]	GBM, SM5.2 R[b]	Experimental[c]
NH_4^+	–79.84	–85.09	–88.6	–79 –81[d]
$CH_3OH_2^+$	–87.40	–77.04	–83.3	–85 –87[d]
$(CH_3)_2OH^+$	–70.47	–63.96	–70.0	–70
$CH_3SH_2^+$	–75.56	–66.41	–73.5	–74
$(CH_3)_2SH^+$	–60.82	–58.64	–66.9	–61
F^-	–107.02	–97.35	–107.8	–107
Cl^-	–76.77	–76.05	–76.9	–77 –78[d]
Br^-	–71.89	–69.76	–71.8	–72
I^-	–64.28	–66.14	–63.1	–63
OH^-	–108.13	–97.77	–109.3	–106 –110[d]
SH^-	–73.00	–74.39	–82.3	–76
CH_3O^-	–94.31	–82.64	–82.3	–95 –98[d]

[a] Reference 95
[b] Reference 131
[c] Reference 170 unless otherwise indicated. The original values have been converted to a 1 molal ideal solution standard state (see reference 176)
[d] Reference 171

Hawkins et al.[131] addressed the hydration of 43 ions by means of the SM5.2R version of the GBM approach, with an MNDO Hamiltonian. On the average, they were within 3.4 kcal/mole of experiment. Table 12 compares some of the PCM and GBM free energies of hydration. In most cases, the two are very similar (when the Barone et al. radii are utilized); for four ions, however, the PCM are considerably better.

Table 13 gives the three contributions to $\Delta G_{hydration}$ for five polyatomic ions, as computed by the PCM technique of Barone et al.,[95] using their radii. As mentioned earlier, $G_{electrostatic}$, $G_{cavitation}$ and G_{vdW} are generally of the same order of magnitude for molecular solutes (Table 8); however this is not true for ionic ones, as can be seen in Table 13. $G_{electrostatic}$ is now often an order of magnitude larger.

(*iv*) *Ionic Solvation via Hybrid Solvent Models*

For ionic as for molecular solutes (Section III.3), some studies have applied the discrete molecular model to the solvent in the immediate environment of the solute, and treated the remainder as a continuum. This can in principle help to deal with the problem of inner-shell structure as well as that of long-range effects. Thus Straatsma and Berendsen used the Born equation to *correct* simulation-obtained free energies of hydration for six monatomic ions.[174] This helped in some instances but not in others.

Table 13
Components of $\Delta G_{hydration}$ For Several Polyatomic Ions, In kcal/mole, As Obtained By A PCM Procedure.[a]

Ion	$G_{electrostatic}$	$G_{cavitation}$	G_{vdW}
H_3O^+	–105.05	3.76	–4.06
NH_4^+	–79.40	4.84	–5.28
CH_3O^-	–94.18	5.56	–5.69
CH_3COO^-	–78.78	9.11	–8.44
$C_6H_5NH_3^+$	–64.80	14.41	–13.69

[a] Reference 95.

A more elaborate hybrid method is that devised by Leontyev et al.,[143,144] which was described in Section III.3. For a group of polyatomic cations, the computed $G_{electrostatic}$ was found to be within an absolute average of 1.5 kcal/mole of the experimental values (the latter were extracted from measured $\Delta G_{solvation}$);[144] for three polyatomic anions, the deviation was 5.7 kcal/mole. (The range was −57.1 to −100.4 kcal/mole.) It was concluded that the anions require separate parametrization.

An interesting combined use of discrete molecular and continuum techniques was demonstrated by Floris et al.[181,182] They used the PCM to develop effective pair potentials and then applied these to molecular dynamics simulations of metal ion hydration. Another approach to such systems is to do an *ab initio* cluster calculation for the first hydration shell, which would typically involve four to eight water molecules, and then to depict the remainder of the solvent as a continuum. This was done by Sanchez Marcos et al. for a group of five cations;[183] the continuum model was that developed by Rivail, Rinaldi et al.[14,108-112] (Section III.2.ii). Their results are compared in Table 14 with those of Floris et al.,[139] who used a similar procedure but PCM-based. In

Table 14
$\Delta G_{hydration}$ For Metal Ions, In Kcal/Mole, Computed By Two Hybrid Discrete Molecular/Continuum Methods

Ion	n^a	$\Delta G_{hydration}$		
		Sanchez Marcos et al.[b]	Floris et al.[c]	Experimental[d]
Li^{+}	4		−115.9	−115.0
Be^{+2}	4	−557.3	−545.7	−574.6
Mg^{+2}	6	−444.8	−439.6	−439.3
Ca^{+2}	6	−342.4	−350.6	
Ca^{+2}	8		−359.0	−362.1
Ca^{+2}	9		−358.2	
Ni^{+2}	6		−459.4	−477.5
Zn^{+2}	6	−427.1	−447.1	−469.2
Al^{+3}	6	−1041.3	−1112.4	−1082.9

[a] *n* is the number of water molecules in the first hydration shell; these are treated explicitly
[b] Reference 183
[c] Reference 139
[d] Reference 178

the latter study, $\Delta G_{hydration}$ was computed for different numbers of water molecules in each ion's first hydration shell to ascertain the most stable configuration; for Ca^{+2}, which is the largest ion in this group,[119] this was found to be $n = 8$, as is shown.

IV. LINEAR SOLVATION ENERGY RELATIONSHIPS (LSER)

1. Empirical

Some years ago, Kamlet and Taft embarked upon a study of how solvents influence the properties of solutes, focusing initially upon the effects of hydrogen bonding upon electronic transitions.[184,185] This led eventually to an empirical relationship between a spectroscopic property X of a given solute, e.g., the position or intensity of a peak, and certain solvent parameters, α, β and π^*:[186]

$$X = X_0 + a\alpha + b\beta + s\pi^* \tag{43}$$

X_0 is the value of the property in the gas phase. (In practice, X and X_0 are often the logarithm of the property in question.) The parameters α and β are measures of a solvent's ability to donate and accept hydrogen bonds, respectively, and π^* is an index of its polarity/polarizability. They were initially assigned on the basis of ultraviolet spectral shifts of certain dyes in a variety of solvents, and hence were labeled *solvatochromic* parameters.[186-188]

The parameters α and β indicate the capacity of a solvent to donate or accept a hydrogen bond from a solute, i.e., the solvent's hydrogen bond *acidity* or *basicity*. π^* is intended to reflect van der Waals-type solute-solvent interactions (dipolar, dispersion, exchange-repulsion, etc.). Equation (43) was subsequently expanded to include a term representing the need to create a cavity for the solute (and thus to interrupt solvent-solvent interactions).[188] For this purpose was used the Hildebrand solubility parameter, δ, which is defined as the square root of the solvent's energy of vaporization per unit volume.[189] Thus Eq. (43) becomes,[190]

$$X = X_0 + a\alpha + b\beta + s\pi^* + h\delta^2 \qquad (44)$$

(In earlier papers, δ entered to the first power.)[188]

While the original applications of the solvatochromic equations, Eqs. (43) and (44), were to correlating ultraviolet, visible and infrared spectral features of solutes, Kamlet, Taft et al. found that they were effective for numerous other properties as well, including reaction rate constants, enthalpies and free energies of solvation, enthalpies and free energies of complex formation, equilibrium constants, NMR shifts and coupling constants, partition coefficients, etc.[188,191] In many instances, only one or two of the parameters in Eqs. (43) and (44) were required. For example, the rate constant for the reaction of tri-*n*-propylamine with methyl iodide in 44 organic solvents can be correlated with π^* alone with R = 0.981.[192] For the $\Delta G_{solvation}$ of *t*-butylchloride, 2-butanone and nitromethane in various organic solvents, expressed in terms of δ and π^*, R = 0.959, 0.924 and 0.991.[193] For the nonpolar solute *n*-octane in aliphatic hydrocarbons, δ suffices to give $\Delta G_{solvation}$ with R = 0.974.[188] Since many of the correlations that have been established reflect free energies in solution, they are known collectively as linear free energy or linear solvation energy relationships (LFER or LSER).

The success of Eqs. (43) and (44) in describing how a given property of a particular solute is affected by different solvents suggested that an analogous approach might be feasible for showing how a property changes from one solute to another in a given solvent. Accordingly, Abraham et al. developed solute parameters that are counterparts to the solvents' α, β and π^*;[194-200] they were designated as α_2^H, β_2^H and π_2^* (sometimes written as π_2^H). These have been incorporated in a number of relationships,[198-201] of the general form,

$$X = \kappa + a\,\alpha_2^H + b\,\beta_2^H + s\,\pi_2^* + rR_2 + vV_x \qquad (45)$$

In Eq. (45), κ is a constant. R_2 is related to the solute's molar refraction[198] and can be viewed as a correction to the polarity/polarizability term, analogous to one that has sometimes

been included in Eqs. (43) and (44).[186-188] V_x is a solute's *characteristic molecular volume*,[202,203] which can readily be calculated from its molecular structure. An equation that combines both solute and solvent parameters has also been used on occasion,[199,201,204] for example,

$$X = \kappa + a\alpha_1\beta_2^H + b\beta_1\alpha_2^H + s\pi_1^*\pi_2^* + vV_x \qquad (46)$$

The subscripts 1 and 2 refer to solvent and solute, respectively.

Relationships such as Eqs. (45) and (46) have been utilized extensively in correlating solubility properties (such as gas/liquid and liquid/liquid partition coefficients), retention volumes in gas/solid chromatography, capacity factors in high-pressure liquid chromatography, etc.[199,200] For instance, gas/liquid partition coefficients for each of 35 different liquid stationary phases were represented with $R \geq 0.985$.[205] Other applications have been in biochemical and pharmacological areas,[199,200] e.g., enzyme inhibition and pollutant effects.

The LSER that have been discussed have had considerable success in correlating a variety of properties, which means that they have, in general, a good predictive capability. They also provide insight, through the magnitudes of the coefficients in Eqs. (43)–(46), into the factors that play important roles in specific types of solute-solvent interactions. Since the parameters are experimentally-derived, however, predictions can be made only for existing compounds. (Hickey and Passino-Reader have suggested rules for estimating the LSER parameters.)[206] There have accordingly been proposed a number of procedures for correlating and predicting solution properties from parameters that can be obtained theoretically or computationally; brief descriptions of some of these have been given by Famini and Wilson[207] and Murray et al.[208] The next Section will focus upon one of these approaches, which is patterned after the LSER of Kamlet, Taft, Abraham et al., Eqs. (43)–(46).

2. Theoretical

Famini, Wilson et al. have correlated a wide array of physical, chemical and biological properties by means of theoretical linear

solvation energy relationships (TLSER).[207,209] These were designed to be theoretical analogs of the empirical LSER, Eqs. (43)–(46). Thus, for a given solute in a series of solvents,

$$X = X_0 + a\varepsilon_{A1} + b\varepsilon_{B1} + c\,q_1^+ + d\,q_1^- + s\,\pi_1^* + h\,\delta_1^2 \qquad (47)$$

while for a series of solutes in one solvent,

$$X = \kappa + a\varepsilon_{A2} + b\varepsilon_{B2} + c\,q_2^+ + d\,q_2^- + s\,\pi_2^* + v\,V_{mc2} \qquad (48)$$

and finally the combined form,

$$X = \kappa + a\varepsilon_{A1}\varepsilon_{B2} + b\varepsilon_{B1}\varepsilon_{A2} + c\,q_1^+\,q_2^- + d\,q_1^-\,q_2^+ + s\,\pi_1^*\,\pi_2^* + h\,\delta_1^2\,V_{mc2} \qquad (49)$$

In Eqs. (47)–(49), the subscripts 1 and 2 are again used to indicate solvent and solute. Hydrogen-bond-donating ability is now expressed in terms of a covalent and an electrostatic contribution, ε_A and q^+; ε_A is the magnitude of the difference between the energies of the lowest unoccupied molecular orbital (LUMO) of the molecule and the highest occupied (HOMO) of water, while q^+ is the atomic charge of the most positive hydrogen. For hydrogen-bond-accepting ability, ε_B is the magnitude of the difference between the molecule's HOMO and water's LUMO, with q^- the absolute value of the most negative atomic charge. (Water is used simply as a reference system.) The polarity/polarizability index, π^*, is related to the molecule's polarizabilty per unit volume, and V_{mc} is its van der Waals volume. δ is again the Hildebrand parameter. The quantum-mechanical calculations are carried out with semiempirical methods.

These TLSER have been used to develop a large number of correlations,[207,209] particularly involving different sorts of pharmacological activities. For example, acetylcholinesterase inhibition by 44 alkylammonium ions correlates with q_2^+, π_2^* and V_{mc2} with R = 0.958.[209] The toxicities of a group of 28 industrial

pollutants can be expressed in terms of ε_{B2}, π_2^* and V_{mc2} with R = 0.992.[207]

V. DISCUSSION AND SUMMARY

This Chapter has outlined several different approaches to the computational determination of solution properties. Two of these address solute-solvent interactions directly, either treating the effects of individual solvent molecules upon the solute explicitly or by means of a reaction field due to a continuum model of the solvent. The other procedures establish correlations between properties of interest and certain features of the solute and/or solvent molecules. There are empirical elements in all of these methods, even the seemingly more rigorous ones, such as the parameters in the molecular dynamics/Monte Carlo intermolecular potentials, Eqs. (16) and (17), or in the continuum model's $G_{cavitation}$ and G_{vdW}, Eqs. (40) and (41), etc.

The correlation-based techniques can achieve as high levels of accuracy as the direct-interaction ones; for example, the GIPF $\Delta G_{hydration}$ have average absolute deviations from experiment, 0.27 kcal/mole (Section II.2.iii), as good or better than the SCRF $\Delta G_{hydration}$ in Table 6. However the GIPF, LSER and TLSER cannot describe the effects of the solute and solvent upon each other, such as polarization of the former and disruption of the structure of the latter.

The GIPF, LSER and TLSER methodologies have in common that they rely on correlations established from experimental databases. They differ fundamentally, however, in the natures of the parameters used for this purpose. Those in the LSER and TLSER were intended specifically to reflect certain aspects of solute-solvent interactions, i.e., hydrogen-bond-donating and –accepting, polarity/polarizability and cavitation. In contrast, the GIPF quantities, Eq. (9), were designed simply to quantitatively characterize, as well as possible, the features of molecular surface electrostatic potentials. While there are good correlations between some of the GIPF and LSER/TLSER parameters,[208] this is a significant conceptual distinction between

them, and between GIPF and quantitative structure-property relationships (QSPR) in general.

If one's interest is solely in predicting a particular solution property, then the GIPF, LSER and TLSER approaches are fast, inexpensive and can be quite accurate, provided that correlations developed from appropriate databases are available. Even if these remain to be established, that can be done in a relatively short time (~ 1 week) and the resulting relationship, if GIPF or TLSER, can be used even for compounds that may not yet have been synthesized. However if a more complete picture of the interaction is desired, then a discrete molecular or continuum (or hybrid/intermediate) procedure should be used. The continuum techniques are faster, less computationally demanding, and, at least for $\Delta G_{solvation}$, appear to be slightly more accurate. For ionic solutes, they may be better able to treat long-range effects. If first-solvation-shell effects are expected to be especially important, or if the disruption of solvent structure is to be characterized, then a discrete molecular solvent model may be preferred. The appeal of the hybrid methods is of course that they can, at least in principle, combine the strengths of both of the others.

ACKNOWLEDGEMENT

We greatly appreciate the support of the Office of Naval Research, Contract No. N00014-99-1-0393 and Project Officer Dr. Judah M. Goldwasser.

REFERENCES

[1] H. Hellmann, *Einfuhrung in die Quantenchemie*, Franz Deuticke, Leipzig, 1937.
[2] R. P. Feynman, *Phys. Rev*. **56** (1939) 340.
[3] J. O. Hirschfelder, C. F. Curtiss and R. B. Bird, *Molecular Theory of Gases and Liquids*, Wiley, New York, 1954.
[4] J. O. Hirschfelder and W. J. Meath, *Adv. Chem. Phys*. **12** (1967) 3.
[5] G. Chalasinski and M. M. Szczesniak, *Chem. Rev*. **100** (2000) 4227.
[6] A. K. Rappe and E. R. Bernstein, *J. Phys. Chem. A* **104** (2000) 6117.
[7] B. Jeziorski, R. Moszynski and K. Szalewicz, *Chem. Rev.* **94** (1994) 1887.
[8] O. Engkvist, P.-O. Astrand and G. Karlstrom, *Chem. Rev.* **100** (2000) 4087.
[9] J. Gao, *Acc. Chem. Res.* **29** (1996) 298.
[10] P. A. Kollman, *Acc. Chem. Res.* **29** (1996) 461.

[11]M. Orozco, C. Alhambra, X. Barril, J. M. Lopez, M. A. Busquets and F. J. Luque, *J. Mol. Mod.* **2** (1996) 1.
[12]J. Tomasi and M. Persico, *Chem. Rev.* **94** (1994) 2027.
[13]C. J. Cramer and D. G. Truhlar, in *Quantitative Treatments of Solute/Solvent Interactions*, Ed. by P. Politzer and J. S. Murray, Elsevier, Amsterdam, 1994, Ch. 2.
[14]J.-L. Rivail and D. Rinaldi, in *Computational Chemistry: Reviews of Current Trends*, Vol. 1, Ed. by J. Leszczynski, World Scientific, Singapore, 1996, Ch. 4.
[15]J. Tomasi and B. Mennucci, in *Encyclopedia of Computational Chemistry*, Vol. 4, Ed. by P.von R. Schleyer, Wiley, New York, 1998, pp. 2547-2560.
[16]C. J. Cramer and D. G. Truhlar, *Chem. Rev.* **99** (1999) 2161.
[17]M. Born, *Z. Physik* **1** (1920) 45.
[18]R. P. Bell, *Trans. Faraday Soc.* **27** (1931) 797.
[19]J. G. Kirkwood, *J. Chem. Phys*. **2** (1934) 351.
[20]L. Onsager, *J. Am. Chem. Soc*. **58** (1936) 1486.
[21]I. Langmuir, *Colloid Symp. Monogr.* **3** (1925) 48.
[22]F. M. Richards, *Annu. Rev. Biophys. Bioeng*. **6** (1977) 151.
[23]P. K. Weiner, R. Langridge, J. M. Blaney, R. Schaefer, P. A. Kollman, *Proc. Natl. Acad. Sci.* **79** (1982) 3754.
[24]M. L. Connolly, *Science*, **221** (1983) 709.
[25]M. M. Francl, R. F. Hout, Jr. and W. J. Hehre, *J. Am. Chem. Soc.* **106** (1984) 563.
[26]R. F. W. Bader, M. T. Carroll, J. R. Cheeseman and C. Chang, *J. Am. Chem. Soc.* **109** (1987) 7968.
[27]P. Politzer and D. G. Truhlar (eds.), *Chemical Applications of Atomic and Molecular Electrostatic Potentials*, Plenum, New York, 1981.
[28]G. Naray-Szabo and G. G. Ferenczy, *Chem. Rev.* **95** (1995) 829.
[29]M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrezewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dappich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. Petersson, P. Y. Aayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rubuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle and J. A. Pople, *Gaussian 98, Revision A.5* (Gaussian, Inc., Pittsburgh, PA, 1998).
[30]J. S. Murray, Z. Peralta-Inga, P. Politzer, K. Ekanayake and P. LeBreton, *Int. J. Quantum Chem*. **83** (2001) 245.
[31]J. S. Murray and P. Politzer, *J. Mol. Struct. (Theochem)* **425** (1998) 107.
[32]J. S. Murray and P. Politzer, in *Encyclopedia of Computational Chemistry*, Vol. 2, Ed. by P. von R. Schleyer, Wiley, New York, 1998, pp. 912-920.
[33]P. Politzer and J. S. Murray, *Trends Chem. Phys.* **7** (1999) 157.
[34]P. Politzer and J. S. Murray, *Fluid Phase Equil.* **185** (2001) 129.
[35]T. Brinck, J. S. Murray and P. Politzer, *Mol. Phys.* **76** (1992) 609.
[36]J. S. Murray, P. Lane and P. Politzer, *Mol. Phys*. **93** (1998) 187.

[37]J. S. Murray, T. Brinck, P. Lane, K. Paulsen and P. Politzer, *J. Mol. Struct.* (*Theochem*) **307** (1994) 55.
[38]J. S. Murray, T. Brinck, M. E. Grice and P. Politzer, *J. Mol. Struct.* (*Theochem*) **256** (1992) 29.
[39]J. Allinger and N. L. Allinger, *Tetrahedron* **2** (1958) 64.
[40]J. S. Murray, Z. Peralta-Inga and P. Politzer, *Int. J. Quantum Chem.* **75** (1999) 267.
[41]J. S. Murray, Z. Peralta-Inga and P. Politzer, *Int. J. Quantum Chem.* **80** (2000) 1216.
[42]P. Politzer and J. S. Murray, in *Reviews in Computational Chemistry*, Vol. 2, Ed. by K. B. Lipkowitz and D. B.Boyd, VCH, New York, 1991, Ch. 7.
[43]P. Politzer and K. C. Daiker, in *The Force Concept in Chemistry*, Ed. by B. M. Deb, Van Nostrand Reinhold, New York, 1981, Ch. 6.
[44]J. S. Murray and P. Politzer, *J. Chem. Res. S*, (1992) 110.
[45]H. Hagelin, J. S. Murray, T. Brinck, M. Berthelot and P. Politzer, *Can. J. Chem.* **73** (1995) 483.
[46]J. S. Murray, F. Abu-Awwad and P. Politzer, *J. Phys. Chem. A* **103** (1999) 1853.
[47]P. Politzer, J. S. Murray and F. Abu-Awwad, *Int. J. Quantum Chem.* **76** (2000) 643.
[48]J. S. Murray, Z. Peralta-Inga, P. Jin and P. Politzer, unpublished work.
[49]C. Hansch, *Acc. Chem. Res.* **26** (1993) 147.
[50]T. Brinck, J. S. Murray and P. Politzer, *J. Org. Chem.* **58** (1993) 7070.
[51]M. J. Kamlet, R. M. Doherty, M. H. Abraham, Y. Marcus and R. W. Taft, *J. Phys. Chem.* **92** (1988) 5244.
[52]G. R. Famini, C. A. Penski and L. Y. Wilson, *J. Phys. Org. Chem.* **5** (1992) 395.
[53]J. S. Murray, T. Brinck and P. Politzer, *J. Phys. Chem.* **97** (1993) 13807.
[54]J. S. Murray, S. G. Gagarin and P. Politzer, *J. Phys. Chem.* **99** (1995) 12081.
[55]P. Politzer, J. S. Murray and S. G. Gagarin, in *Recent Advances in the Chemistry and Physics of Fullerenes and Related Materials*, Vol. 3, Ed. by K. M. Kadish and R. S. Ruoff, The Electrochemical Society, Pennington, NJ, 1996, pp. 45-52.
[56]P. Politzer, J. S. Murray and P. Flodmark, *J. Phys. Chem.* **100** (1996) 5538.
[57]P. Politzer, P. Lane, J. S. Murray and T. Brinck, *J. Phys. Chem.* **96** (1992) 7938.
[58]J. S. Murray, P. Lane, T. Brinck and P. Politzer, *J. Phys. Chem.* **97** (1993) 5144.
[59]P. Politzer, J. S. Murray, P. Lane and T. Brinck, *J. Phys. Chem.* **97** (1993) 729.
[60]P. Politzer, J. S. Murray, M. C. Concha and T. Brinck, *J. Mol. Struct. (Theochem)* **281** (1993) 107.
[61]M. R. Gholami and B. A. Talebi, *J. Phys. Org. Chem.* **16** (2003) 369.
[62]M. P. Allen and D. J. Tildesley, *Computer Simulations of Liquids*, Clarendon Press, Oxford, 1987.
[63]J. M. Haile, *Molecular Dynamics Simulations*, Wiley, New York, 1992.
[64]W. L. Jorgensen, in *Encyclopedia of Computational Chemistry*, Vol. 3, Ed. by P. von R. Schleyer, Wiley, 1998, pp. 1754-1763.
[65]J. R. Maple, in *Encyclopedia of Computational Chemistry*, Vol. 2, Ed. by P. von R. Schleyer, Wiley, 1998, pp. 1015-1024.
[66]P. E. Smith and B. M. Pettitt, *J. Phys. Chem.* **98** (1994) 9700.
[67]P. Politzer and S. Boyd, *Struct. Chem.* **13** (2002) 105.
[68]M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *J. Am. Chem. Soc.* **107** (1985) 3902.
[69]J. J. P. Stewart, *J. Comput. Chem.* **10** (1989) 209, 221.

[70]J. Aqvist and A. Warshel, *Chem. Rev.* **93** (1993) 2523.
[71]W. J. Hehre, L. Radom, P. von R. Schleyer and J. A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986.
[72]W. F. van Gunsteren, in *Computer Simulation of Biomolecular Systems*, Ed. by W. F. van Gunsteren and P. K. Weiner, ESCOM, Leiden (The Netherlands), 1989, pp. 27-59.
[73]P. A. Kollman and K. M. Merz, Jr., *Acc. Chem. Res.* **23** (1990) 246.
[74]W. L. Jorgensen, in *Encyclopedia of Computational Chemistry*, Vol. 2, Ed. by P. von R.Schleyer, Wiley, 1998, pp. 1061-1070.
[75]P.-Y. Morgantini and P. A. Kollman, *J. Am. Chem. Soc.* **117** (1995) 6057.
[76]W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.* **117** (1995) 5179.
[77]W. L. Jorgensen and J. M. Briggs, *J. Am. Chem. Soc.* **111** (1989) 4190.
[78]P. Cieplak, P. Bash, S. C. Singh and P. A. Kollman, *J. Am. Chem. Soc.* **109** (1987) 6283.
[79]E. C. Meng, J. W. Caldwell and P. A. Kollman, *J. Phys. Chem.* **100** (1996) 2367.
[80]H. A. Carlson, T. B. Nguyen, M. Orozco and W. L. Jorgensen, *J. Comput. Chem.* **14** (1993) 1240.
[81]R. S. Mulliken, *J. Chem. Phys.* **23** (1955) 1833.
[82]W. L. Jorgensen, *Acc. Chem. Res.* **22** (1989) 184.
[83]C. A. Reynolds, *J. Am. Chem. Soc.* **112** (1990) 7545.
[84]J. Gao and X. Xia, *Science* **258** (1992) 631.
[85]J. Gao, F. J. Luque and M. Orozco, *J. Chem. Phys.* **98** (1993) 2975.
[86]T. J. Marrone, D. S. Hartsough and K. M. Merz, Jr., *J. Phys. Chem.* **98** (1994) 1341.
[87]W. R. Smythe, *Static and Dynamic Electricity*, McGraw-Hill, New York, 1950.
[88]A. A. Rashin and B. Honig, *J. Phys. Chem.* **89** (1985) 5588.
[89]*Chemical & Engineering News*, Nov. 17 (2003) 70.
[90]S. Glassstone, *Text-book of Physical Chemistry*, Van Nostrand, New York, 1940.
[91]T. G. Scholte, *Physica* **15** (1949) 437.
[92]A. D. Buckingham, *Trans. Faraday Soc.* **49** (1953) 881.
[93]F. H. Westheimer and J. G. Kirkwood, *J. Chem. Phys*. **6** (1938) 513.
[94]M. J. Huron and P. Claverie, *J. Phys. Chem.* **76** (1972) 2123; **78** (1974) 1853, 1862.
[95]V. Barone, M. Cossi and J. Tomasi, *J. Chem. Phys*. **107** (1997) 3210.
[96]J.-L. Rivail and B. Terryn, *J. Chim. Phys. Physicochim. Biol.* **79** (1982) 1.
[97]D. Rinaldi, *Comput. & Chem.* **6** (1982) 155.
[98]J. Tomasi, R. Bonaccorsi, R. Cammi and F. J. Olivares del Valle, *J. Mol. Struct. (Theochem)* **234** (1991) 401.
[99]G. J. Hoijting, E. de Boer, P. H. van der Meij and W. P. Weijland, *Rec. Trav. Chim. Pays-Bas* **75** (1956) 487.
[100]O. Chalvet and I. Jano, *Compt. Rend. Seances Acad. Sci*. **259** (1964) 1867.
[101]W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.* **112** (1990) 6127.
[102]W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.* **110** (1988) 1657.
[103]S. Miertus and J. Tomasi, *Chem. Phys.* **65** (1982) 239.

[104]A. A. Rashin and K. Namboodiri, *J. Phys. Chem*. **91** (1987) 6003.
[105]F. J. Luque, J. M. Bofill and M. Orozco, *J. Chem. Phys*. **103** (1995) 10183. See also J. G. Angyan, *J Chem. Phys*. **107** 1291 (1997) 1291 and F. J. Luque, J. M. Bofill and M. Orozco, *J. Chem. Phys*. **107** (1997) 1293.
[106]I. N. Levine, *Quantum Chemistry*, 5th ed., Prentice Hall, Upper Saddle River, NJ, 2000.
[107]C. Curutchet, C. J. Cramer, D. G. Truhlar, M. F. Ruiz-Lopez, D. Rinaldi, M. Orozco and F. J. Luque, *J. Comput. Chem*. **24** (2003) 284.
[108]D. Rinaldi and J.-L. Rivail, *Theor. Chim. Acta* **32** (1973) 57.
[109]J.-L. Rivail and D. Rinaldi, *Chem. Phys*. **18** (1976) 233.
[110]D. Rinaldi, M. F. Ruiz-Lopez and J.-L. Rivail, *J. Chem. Phys*. **78** (1983) 834.
[111]V. Dillet, D. Rinaldi and J.-L. Rivail, *J. Phys. Chem*. **98** (1994) 5034.
[112]V. Dillet, D. Rinaldi, J. Bertran and J.-L. Rivail, *J. Chem. Phys*. **104** (1996) 9437.
[113]O. Tapia, in *Quantum Theory of Chemical Reactions*, Vol. II, Ed. by R. Daudel, A. Pullman, L.Salem and A. Veillard, Reidel, Dordrecht, 1981, p. 25.
[114]R. Constanciel and R. Contreras, *Theor. Chim. Acta* **65** (1984) 1.
[115]T. Kozaki, K. Morihashi and O. Kikuchi, *J. Mol. Struct.* **168** (1988) 265.
[116]J. W. Storer, D. J. Giesen, C. J. Cramer and D. G. Truhlar, *J. Comput.-Aided Mol. Design* **9** (1995) 87.
[117]S. Miertus, E. Scrocco and J. Tomasi, *Chem. Phys.* **55** (1981) 117.
[118]R. Montagnani and J. Tomasi, *J. Mol. Struct. (Theochem)* **279** (1993) 131.
[119]D. R. Lide (ed.), *Handbook of Chemistry and Physics*, 78th ed., CRC Press, New.York, 1997.
[120]H. H. Uhlig, *J. Phys. Chem*. **41** (1937) 1215.
[121]R. B. Hermann, *J. Phys. Chem*. **76** (1972) 2754.
[122]D. Sitkoff, K. A. Sharp and B. Honig, *J. Phys. Chem*. **98** (1994) 1978.
[123]C. J. Cramer and D. G. Truhlar, *J. Am. Chem. Soc*. **113** (1991) 8305.
[124]I. Tunon, M. F. Ruiz-Lopez, D. Rinaldi and J. Bertran, *J. Comput. Chem*. **17** (1996) 148.
[125]M. Orozco, W. L. Jorgensen and F. J. Luque, *J. Comput. Chem.* **14** (1993) 1498.
[126]V. Gogonea and E. Osawa, *J. Mol. Struct. (Theochem)* **311** (1994) 305.
[127]M. Zacharias, *J. Phys. Chem. A* **107** (2003) 3000.
[128]R. A. Pierotti, *Chem. Rev*. **76** (1976) 717; P. Claverie, *Perspect. Quantum Chem. Biochem.* **2** (1978) 69.
[129]F. M. Floris and J. Tomasi, *J. Comput. Chem.* **10** (1989) 616.
[130]F. M. Floris, A. Tani and J. Tomasi, *Chem. Phys.* **169** (1993) 11.
[131]G. D. Hawkins, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B* **102** (1998) 3257.
[132]M.-H. Baik and R. A. Friesner, *J. Phys. Chem. A* **106** (2002) 7407.
[133]G. A. A. Saracino, R. Improta and V. Barone, *Chem. Phys. Lett.* **373** (2003) 411.
[134]D. F. Green and B. Tidor, *J. Phys. Chem. B* **107** (2003) 10261.
[135]K. N. Rankin, T. Sulea and E. O. Purisima, *J. Comput. Chem.* **24** (2003) 954.
[136]E. Sanchez Marcos, B. Terryn and J.-L. Rivail, *J. Phys. Chem.* **89** (1985) 4695.
[137]J. Zeng, J. S. Cram, N. S. Hush and J. R. Reimers, *Chem. Phys. Lett.* **206** (1993) 323.

[138]E. Sanchez Marcos, R. R. Pappalardo and D. Rinaldi, *J. Phys. Chem.* **95** (1991) 8928.
[139]F. M. Floris, M. Persico, A. Tani and J. Tomasi, *Chem. Phys.* **195** (1995) 207.
[140]C. Aleman and S. E. Galembeck, *Chem. Phys.* **232** (1998) 151.
[141]M. Brauer, M. Mosquera, J. L. Perez-Lustres and F. Rodriguez-Prieto, *J. Phys. Chem. A* **102** (1998) 10736.
[142]C. Aleman, *Chem. Phys. Lett.* **302** (1999) 461.
[143]I. V. Leontyev, M. V. Vener, I. V. Rostov, M. V. Basilevsky and M. D. Newton, *J. Chem. Phys.* **119** (2003) 8024.
[144]M. V. Vener, I. V. Leontyev and M. V. Basilevsky, *J. Chem. Phys.* **119** (2003) 8038.
[145]P. F. B. Goncalves and H. Stassen, *J. Comput. Chem.* **23** (2002) 706; **24** (2003) 1758.
[146]A. Warshel and M. Levitt, *J. Mol. Biol*. **103** (1976) 227.
[147]A. Warshel and S. T. Russell, *Quantum Rev. Biophys.* **17** (1984) 285.
[148]A. Warshel and Z. T. Chu, in *Structure and Reactivity in Aqueous Solution. Characterization of Chemical and Biological Systems*, Ed. by C. J. Cramer and D. G. Truhlar, ACS Symposium Series 568, American Chemical Society, Washington, DC, 1994, pp. 71-94.
[149]P. Langevin, *Compt. Rend. Acad. Sci. Paris* **146** (1908) 530.
[150]J. Warwicker and H. C. Watson, *J. Mol. Biol.* **157** (1982) 671.
[151]D. T. Edmonds, N. K. Rogers and M. J. E. Sternberg, *Mol. Phys.* **52** (1984) 1487.
[152]M. K. Gilson, K. A. Sharp and B. H. Honig, *J. Comput. Chem*. **9** (1988) 327.
[153]M. E. Davis and J. A. McCammon, *J. Comput. Chem*. **10** (1989) 387; **12** (1991) 909.
[154]W. H. Orttung, *J. Am. Chem. Soc*. **100** (1978) 4369.
[155]R. Luo, J. Moult and M. K. Gilson, *J. Phys. Chem. B* **101** (1997) 11226.
[156]S. A. Hassan, F. Guarnieri and E. L. Mehler, *J. Phys. Chem. B* **104** (2000) 6478, 6490.
[157]A. Jean-Charles, A. Nicholls, K. Sharp, B. Honig, A. Tempczyk, T. F. Hendrickson and W. C. Still, *J. Am. Chem. Soc.* **113** (1991) 1454.
[158]R. D. Coalson and T. L. Beck, in *Encyclopedia of Computational Chemistry*, Vol. 3, Ed. by P. von R. Schleyer, Wiley, New York, 1998, pp. 2086-2100.
[159]B. Honig, K. Sharp and A.-S. Yang, *J. Phys. Chem.* **97** (1993) 1101.
[160]N. G. J. Richards, in *Molecular Orbital Calculations for Biological Systems*, Ed. by A.-M. Sapse, Oxford University Press, New York, 1998, Ch. 2.
[161]D. Bashford, in *Encyclopedia of Computational Chemistry*, Vol. 3, Ed. by P. von R. Schleyer, Wiley, New York, 1998, pp. 1542-1548.
[162]B. E. Conway and J. O'M. Bockris, in *Modern Aspects of Electrochemistry*, vol. 1, Ed. by J. O'M. Bockris, Butterworths, London, 1954, Ch. 2.
[163]J. I. Padova, in *Modern Aspects of Electrochemistry*, Vol. 7, Ed. by B. E. Conway and J. O'M. Bockris, Plenum Press, New York, 1972, Ch. 1.
[164]J. D. Bernal and R. H. Fowler, *J. Chem. Phys.* **1** (1933) 515.
[165]W. M. Latimer, K. S. Pitzer and C. M. Slansky, *J. Chem. Phys.* **7** (1939) 108.
[166]E. J. W. Verwey, *Rec. Trav. Chim*. **61** (1942) 127, 564.
[167]K. P. Mischenko, *Acta Physicochim.* URSS **3** (1935) 693.
[168]D. R. Rosseinsky, *Chem. Rev.* **65** (1965) 467.
[169]J. Burgess, *Metal Ions in Solution*, Ellis Horwood, Chichester, UK, 1978.

[170]R. G. Pearson, *J. Am Chem. Soc.* **108** (1986) 6109.
[171]J. Florian and A. Warshel, *J. Phys. Chem. B* **101** (1997) 5583.
[172]M. Mezei and D. L. Beveridge, *J. Chem. Phys.* **74** (1981) 6902.
[173]J. Chandrasekhar, D. C. Spellmeyer and W. L. Jorgensen, *J. Am. Chem. Soc.* **106** (1984) 903.
[174]T. P. Straatsma and H. J. C. Berendsen, *J. Chem. Phys.* **89** (1988) 5876.
[175]B. Jayaram, R. Fine, K. Sharp and B. Honig, *J. Phys. Chem.* **93** (1989) 4320.
[176]P. B. Balbuena, K. P. Johnston and P. J. Rossky, *J. Phys. Chem.* **100** (1996) 2706.
[177]J. O'M. Bockris and A. K. N. Reddy, *Modern Electrochemistry*, Vol. 1, Plenum Press, New York, 1970, Ch. 2.
[178]Y. Marcus, *Ion Solvation*, Wiley, Chichester, UK, 1985.
[179]W. L. Jorgensen, J. F. Blake and J. K. Buckner, *Chem. Phys.* **129** (1989) 193.
[180]R. M. Noyes, *J. Am. Chem. Soc.* **84** (1962) 513.
[181]F. Floris, M. Persico, A. Tani and J. Tomasi, *Chem. Phys. Lett.* **199** (1992) 518.
[182]F. Floris, M. Persico, A. Tani and J. Tomasi, *Chem. Phys. Lett.* **227** (1994) 126.
[183]E. Sanchez Marcos, R. R. Pappalardo and D. Rinaldi, *J. Phys. Chem.* **95** (1991) 8928.
[184]M. J. Kamlet and R. W. Taft, *J. Am. Chem. Soc.* **98** (1976) 377.
[185]R. W. Taft and M. J. Kamlet, *J. Am. Chem. Soc.* **98** (1976) 2886.
[186]M. J. Kamlet, J.-L. M. Abboud and R. W. Taft, *J. Am. Chem. Soc.* **99** (1977) 6027.
[187]J.-L. M. Abboud, M. J. Kamlet and R. W. Taft, *J. Am. Chem. Soc.* **99** (1977) 8325.
[188]M. J. Kamlet, J.-L. M. Abboud, M. H. Abraham and R. W. Taft, *J. Org. Chem.* **48** (1983) 2877.
[189]A. F. M. Barton, *Chem. Rev.* **75** (1975) 731.
[190]M. J. Kamlet, R. M. Doherty, G. R. Famini and R. W. Taft, *Acta Chem. Scand. B* **41** (1987) 589.
[191]M. J. Kamlet, J.-L. M. Abboud and R. W. Taft, *Prog. Phys. Org. Chem.* **13** (1981) 485.
[192]M. J. Kamlet, T. N. Hall, J. Boykin and R. W. Taft, *J. Org. Chem.* **44** (1979) 2599.
[193]M. H. Abraham, M. J. Kamlet and R. W. Taft, *J. Chem. Soc., Perkin Trans. II*, (1982) 923.
[194]M. H. Abraham, P. P. Duce, P. L. Grellier, D. V. Prior, J. J. Morris and P. J. Taylor, *Tetrahedron Lett.* **29** (1988) 1587.
[195]M. H. Abraham, P. L. Grellier, D. V. Prior, J. J. Morris, P. J. Taylor, C. Laurence and M. Berthelot, *Tetrahedron Lett*. **30** (1989) 2571.
[196]M. H. Abraham, P. L. Grellier, D. V. Prior, P. P. Duce, J. J. Morris and P. J. Taylor, *J. Chem. Soc., Perkin Trans. II*, (1989) 699.
[197]M. H. Abraham, P. L. Grellier, D. V. Prior, J. J. Morris and P. J. Taylor, *J. Chem. Soc. Perkin Trans. II*, 521 (1990).
[198]M. H. Abraham, G. S. Whiting, R. M. Doherty and W. J. Shuely, *J. Chem. Soc., Perkin Trans. II*, 1451 (1990).
[199]R. W. Taft and J. S. Murray, in *Quantitative Treatments of Solute/Solvent Interactions*, Ed. by P. Politzer and J. S. Murray, Elsevier, Amsterdam, 1994, Ch. 3.

[200]M. H. Abraham, in *Quantitative Treatments of Solute/Solvent Interactions*, Ed. by P. Politzer and J. S. Murray, Elsevier, Amsterdam, 1994, Ch. 4.
[201]M. H. Abraham, G. S. Whiting, R. Fuchs and E. J. Chambers, *J. Chem. Soc. Perkin Trans. II*, (1990) 291.
[202]J. C. McGowan, *J. Appl. Chem. Biotechnol*. **28**, 599 (1978); **34A** (1984) 38.
[203]M. H. Abraham and J. C. McGowan, *Chromatographia* **23** (1987) 243.
[204]M. J. Kamlet, M. H. Abraham, R. M. Doherty and R. W. Taft, *J. Am. Chem. Soc.* **106** (1984) 464.
[205]M. H. Abraham and G. S. Whiting, *J. Chromatogr*. **594** (1992) 229.
[206]J. P. Hickey and D. R. Passino-Reader, *Environ. Sci. Technol*. **25** (1991) 1753.
[207]G. R. Famini and L. Y. Wilson, in *Quantitative Treatments of Solute/Solvent Interactions*, Ed. by P. Politzer and J. S. Murray, Elsevier, Amsterdam, 1994, Ch. 7.
[208]J. S. Murray, P. Politzer and G. R. Famini, *J. Mol. Struct. (Theochem)* **454** (1998) 299.
[209]C. J. Cramer, G. R. Famini and A. H. Lowrey, *Acc. Chem. Res*. **26** (1993) 599.

2

Porous Silicon: Morphology and Formation Mechanisms

Gregory X. Zhang

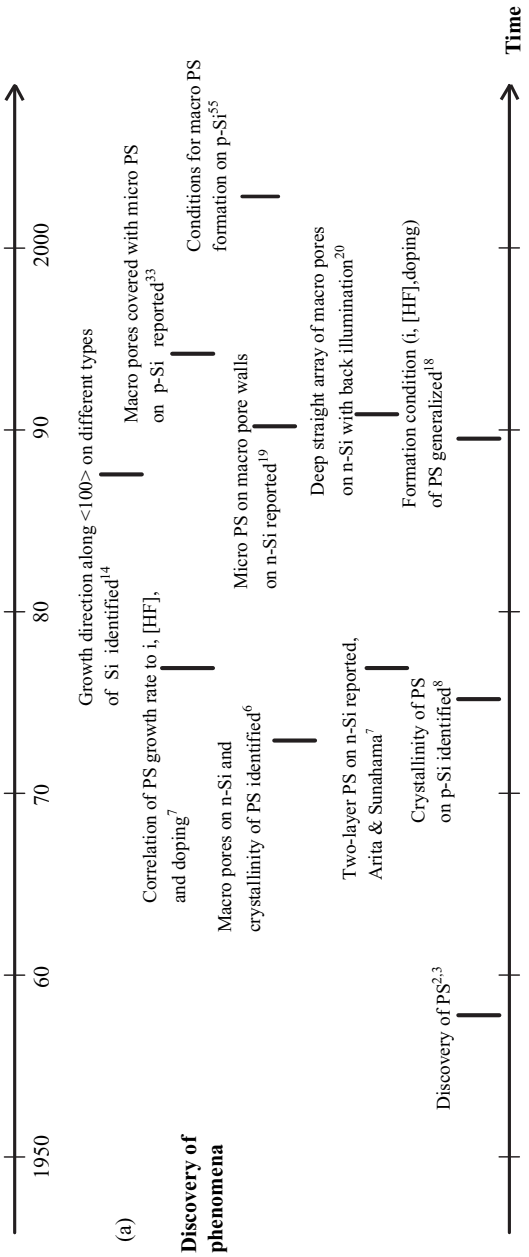
Teck Cominco, Product Technology Centre, Mississauga, Ontario, Canada

I. INTRODUCTION

Porous silicon (PS) is a material formed by anodic dissolution of single crystalline silicon in HF containing solutions. Since its discovery more than four decades ago, a large number of investigations have been undertaken, the results of which revealed that PS has extremely rich morphological features and the formation process of PS is a very complex function of numerous factors.[1] Accordingly, many theories have been proposed on the various mechanistic aspects on formation and morphology of PS. Figure 1 is a summary of the progress of research on PS with respect to the discovery of major PS features and development of theories.

Uhlir and Turner[2,3] first reported in the late 1950's that silicon surface can be covered with a brown film during anodization in HF solutions and suggested that the film was a subfluoride $(SiF_2)_x$ grown on the surface during the anodic dissolution. Later, Memming and Schwandt proposed that the

Modern Aspects of Electrochemistry, Number 39, edited by C. Vayenas et al., Springer, New York, 2005.


(a)
Discovery of phenomena
1950
60
70
80
90
2000
Time
1950
Discovery of PS[2,3]
Crystallinity of PS on p-Si identified[8]
Two-layer PS on n-Si reported, Arita & Sunahama[7]
Macro pores on n-Si and crystallinity of PS identified[6]
Correlation of PS growth rate to i, [HF], and doping[7]
Growth direction along <100> on different types of Si identified[14]
Micro PS on macro pore walls on n-Si reported[19]
Deep straight array of macro pores on n-Si with back illumination[20]
Formation condition (i, [HF].doping) of PS generalized[18]
Macro pores covered with micro PS on p-Si reported[33]
Conditions for macro PS formation on p-Si[55]

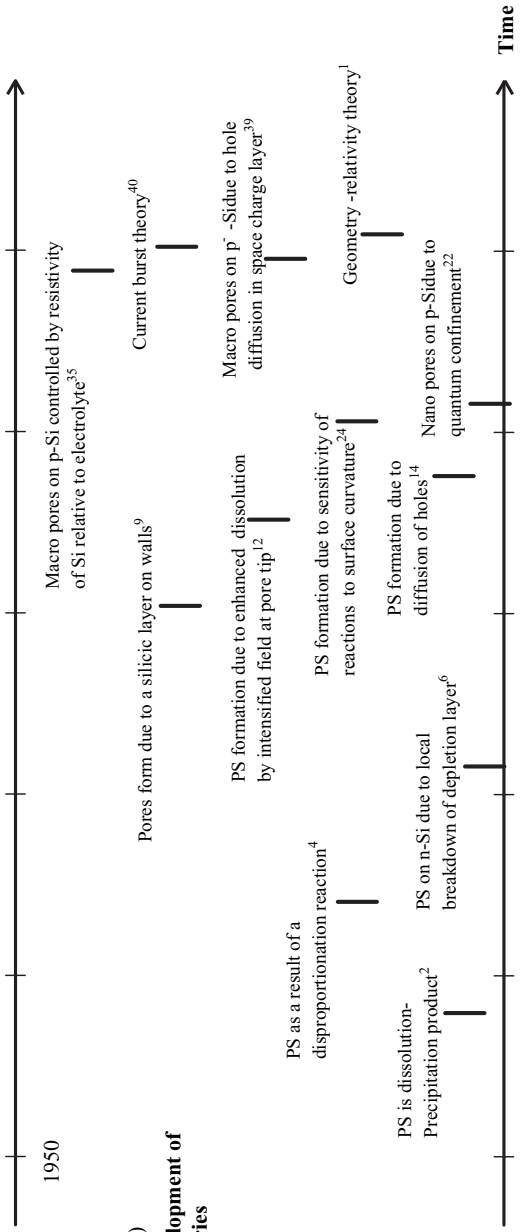


Figure 1. Progress in the discovery of morphological phenomena and in the development of theories on formation mechanisms of porous silicon.

brown film was a dissolution/precipitation product ($Si^0_{amorphous}$) resulted from a two-step disproportionation reaction.[4]

The formation of etch pits and tunnels on n-Si during anodization in HF solutions was reported in the early 1970's. It was found that the solid surface layer is the remaining substrate silicon left after anodic dissolution. The large current observed on n-Si at an anodic potential was postulated to be due to barrier breakdown.[5,6] By early 80's[7-11] it was established that the brown films formed by anodization on silicon substrate of all types are a porous material with the same single crystalline structure as the substrate.

Many theories on the formation mechanisms of PS emerged since then. Beale et al.[12] proposed that the material in the PS is depleted of carriers and the presence of a depletion layer is responsible for current localization at pore tips where the field is intensified. Smith et al.[13-15] described the morphology of PS based on the hypothesis that the rate of pore growth is limited by diffusion of holes to the growing pore tip. Unagami[16] postulated that the formation of PS is promoted by the deposition of a passive silicic acid on the pore walls resulting in the preferential dissolution at the pore tips. Alternatively, Parkhutik et al.[17] suggested that a passive film composed of silicon fluoride and silicon oxide is between PS and silicon substrate and that the formation of PS is similar to that of porous alumina.

The conditions for formation of PS on all types of substrates in terms of current density and HF concentrations were established by Zhang et al.[18] by the end of 1980's. Whether formation of PS occurs during anodization was found to be largely independent of the electronic properties of silicon such as doping type and concentration, but depend on the nature of electrochemical reactions.

Lehmann and Foll[9-11] in early 1990's reported the formation of straight, smooth and well spaced macro pore arrays on n-Si using backside illumination and surface patterning. They postulated that for these large straight pores the dissolution rate at pore tips are limited by mass transfer in the electrolyte and that the relative rates of carrier transport in the silicon semiconductor and mass transport in the electrolyte determine PS morphology. For the micro pores formed on p-Si, Lehmann and Gosele[22-23] proposed a quantum confinement model; due to the quantum

confinement the pore walls are depleted of carriers and thus do not dissolve during the anodization. At about the same time, Zhang[24] proposed the surface curvature model, which postulated that not only the rate but also the distribution of reactions on a curved pore bottom can be greatly affected by surface curvature.

In the late 1990's, Foll et al.[40-43] proposed the current burst theory. The basic hypothesis is that the electrochemical reactions involved in the dissolution of the silicon surface operate in microscopic units. These reaction units have a temporal and a spatial distribution in number and in the state of activity. The formation of pores is due to the synchronization of these operation units at certain time and geometrical scales.

Two-layer PS with a micro PS on surface of a macro PS formed on n-Si under front illumination had been reported in the late 70's but was little investigated until the 90's.[26-29] The micro PS may have a fractal-like geometry; the structure can vary in the same layer from amorphous to single crystalline and diameter from a few nm to hundreds of nm. It appeared that while the formation of a macro PS layer under front illumination follows the same mechanism as the macro PS formed in the dark, the formation of micro PS is due to the effect of the photo generated carriers. According to Arita[30] the drift current due to the photo carriers generated in the depletion layer is responsible for the amorphous-like PS. Alternatively, Clement et al.[31] suggested that the micro PS formed under illumination could result from shattering of the macro PS into fine filaments due to residual stress.

Formation of macro PS and two-layer PS on lowly doped p-Si in anhydrous organic HF solutions was first reported by Propst and Kohl[32,33] in the mid 1990's. It was thought to be related to the chemistry of the organic solvents. Macro pores on lowly doped p-Si were later found to also form in aqueous solutions. Walls of the macro pores are not always covered by a micro PS layer.[34-36] Wehrspohn et al.[37,38] suggested that a necessary condition for the formation of macro PS on lowly doped p-Si is a higher resistivity of the substrate than that of electrolyte. This idea was soon invalidated as macro PS was also found to occur in electrolyses that have much higher resistance than the silicon substrate.[34,39] Alternatively, Lehmann and Ronnebeck[39] postulated that the formation of macro PS on lowly doped p-Si is due to the

dominant effect of thermionic emission which is sensitive to barrier height rather than barrier width. More recently, Zhang[1] proposed that the formation of macro PS on lowly doped p-Si is due to the non-linear potential distribution in the highly resistive substrate caused by surface curvature.

Some ideas have been elaborated with extensive mathematical formulation. For example, a theoretical modeling based on charge transfer kinetics for PS morphology was attempted by Jaguiro et al.[44] Similar mathematical modeling considering the transport phenomena of carriers in the semiconductor, ions in the electrolyte and surface tension was proposed by Valance.[45] An analytical formulation based principally on thermodynamical arguments was offered by Monica et al.[46]

Most theories have dealt only with certain aspects of the very complex morphology and formation of PS. No single theory is yet close to a globally coherent description of the extremely rich and complex nature of the reactions and the resulting morphological evolution of the silicon electrode surface involved in the formation of PS.[1] Also, these theories on PS did not take into account for the many details of anodic electrochemical reactions and processes of the silicon electrolyte interface, which had been extensively investigated. However, a complete description of the diverse morphological features of porous silicon requires the integration of the all the morphological aspects as well as the fundamental electrochemical reaction processes involved in silicon/electrolyte interface. Any theory for the mechanisms of porous silicon formation without such a global integration will be limited in the scope of its validity and the power to explain morphological details. In addition, a globally and microscopically accurate model also requires the full characterization of all the morphological features of PS.

A detailed and comprehensive review on all aspects of the fundamental and applied electrochemistry of silicon/electrolyte interface was provided in a recently published book.[1] The objective of this paper is to provide a conceptual analysis of the mechanisms for the morphology and formation of porous silicon using the large body of the information assembled in the book and to provide an integrated view of the formation mechanisms that can be coherent with the various morphological features on the

one hand and with the understanding of the fundamental anodic reaction processes on the other.

II. FORMATION OF POROUS SILICON

1. I-V Curves and Formation Condition

The formation condition for PS can be best characterized by i-V curves. Figure 2 shows a typical i-V curve of silicon in a HF solution.[56] At small anodic overpotentials the current increases exponentially with electrode potential. As the potential is increased, the current exhibits a peak and then remains at a relatively constant value. At potentials more positive than the current peak the surface is completely covered with an oxide film and the anodic reaction proceeds through the formation and dissolution of oxide, the rate of which depends strongly on HF concentration. Hydrogen evolution simultaneously occurs in the exponential region and its rate decreases with potential and almost ceases above the peak value.

Porous silicon forms in the exponential region but not at potentials higher than that corresponding to the current peak, J_1. Electrochemical polishing occurs at potentials higher than the peak potential. At the potentials above the exponential region and below the potential of the current peak, PS formation also occurs, but the PS layer does not completely cover the sample surface.[18] The surface coverage of PS decreases as the potential approaches the peak value. The different regions are similar for other fluoride containing solutions of various compositions and pH values.[4,52,57,58]

The characteristics of the i-V curves are largely identical for all types materials except for non-heavily doped.[5,18] For non-heavily doped n-Si, it requires a large polarization or illumination to generate anodic reaction as the surface is under a reverse bias as shown in Figure 3.[18] The i-V curves may not show the clearly defined regions as on other types of Si with a positive potential scanning rate. Different regions can be revealed on the curve measured at certain negative scanning rate. Also, the current density on n-Si is very sensitive to surface roughness.[6,51]

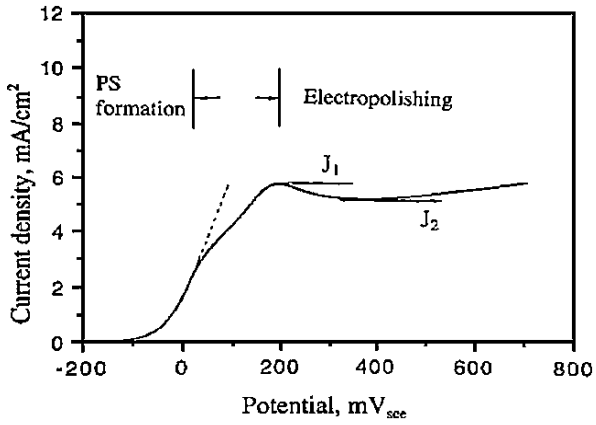


Figure 2. Typical anodic i-V curve measured on a moderately doped p-Si in 1% HF solution.[18]

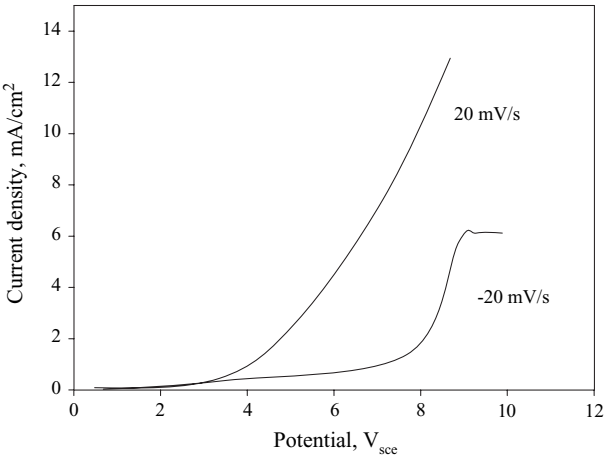


Figure 3. i-V curves of n-Si in 1% HF solution.[18]

Mechanical roughening can greatly increase the current density at a given potential. Current can arise on a rough surface at a potential several volts lower than that on a smooth surface.

Anodic dissolution of n-Si can also proceed at a polarization under illumination. The maximum current is limited by illumination intensity when the saturation photo current density is lower than the critical current, J_1. The characteristics of i-V curves of n-Si under a high illumination intensity, when the reaction is no longer limited by the availability of photo generated carriers, is identical to that for p-Si. Similar also to p-Si, formation of PS on n-Si occurs only below the critical current, J_1.[24]

The dependence of i-V characteristics on doping is an indication of the carrier conduction mechanisms that can occur in a silicon substrate. When carrier supply in the silicon substrate is not rate limiting, i-V curves obtained in different HF concentrations are identical in characteristics.[56] When the supply of carriers in the substrate is rate limiting as determined by the doping and illumination condition, i-V curves display distinct difference between p-Si and n-Si, and for n-Si between illuminated and non-illuminated materials. For non-heavily doped p-Si, the process is by thermal emission of holes while it is by Zener tunnelling for heavily doped materials.[10,59] For moderately doped n-Si the reaction is limited by the minority holes; significant current occurs when a large amount of holes are generated by illumination or when relatively high anodic potentials are applied to allow interface tunneling to occur.[11,62]

Examination of the sample surfaces that are anodized at different potentials indicates that the potential corresponding to the maximum slope of the i-V curve is the upper limit for formation of a porous silicon layer that covers uniformly the entire surface. At potentials between the maximum slope and current peak, porous layer may still form but the surface coverage is not uniform. Plotting the current at the maximum slope and the peak current for different types of silicon as a function of HF concentration, the condition for occurrence of PS formation and electropolishing is obtained as shown in Figure 4.[18] This figure also shows that the three regions in relation to current density and HF concentration are essentially independent of the silicon substrate doping type and concentration. This means that the

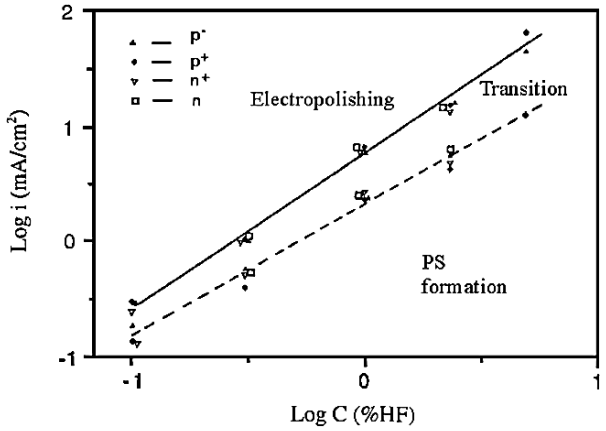


Figure 4. Formation condition for porous silicon; solid line - peak current density, dotted line - current density at the maximum slope (see Figure.2).[18]

differences in semiconducting properties of the silicon samples have little effect on the occurrence of these regions. The various parameters involved in PS formation such as potential, doping, and illumination, affect the occurrence of different regions through their relation to the current density. Low current density and high HF concentration favour PS formation while high current and low HF concentration favours polishing.

Electropolishing region does not occur in anhydrous organic solutions due to the lack of water which is required for the formation of oxide film. Figure 5, as an example, shows that in anhydrous HF-MeCN solution the current can increase with potential to a value of about 0.5 A/cm^2 without showing a peak current. The relationship between current and potential is linear due to the rate limiting effect of resistance in solution and silicon substrate.

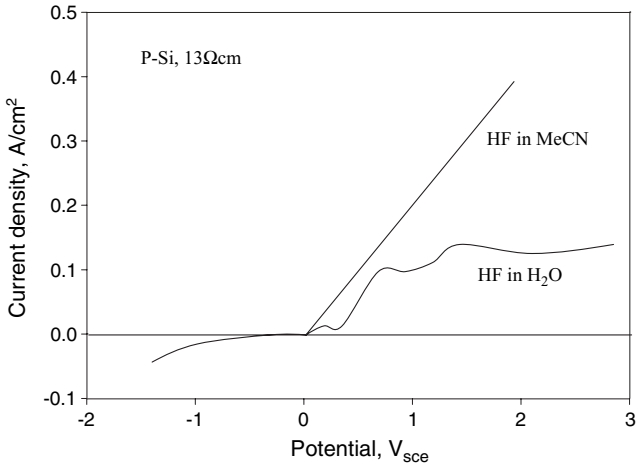


Figure 5. i-V behavior of p-Si in 2M HF + 0.25M TBAP in MeCN and in 2M HF + 1M NH_4F in H_2O, After Propst and Kohl.[33]

2. Hydrogen Evolution and Effective Dissolution Valence

Evolution of hydrogen gas occurs during the formation of PS in aqueous solutions. The amount of hydrogen gas is proportional to the time of anodization.[4] Due to hydrogen evolution, which is a chemical reaction, the effective dissolution valence of silicon, n, can vary between 2 to 4 depending on the relative contribution of this reaction. In general, n increases with increasing anodic current density as the amount of hydrogen evolution reduces; the change is sharpest near the peak current density J_1. At a given current density, n increases with HF concentration. Also, dissolution valence decreases with increasing PS layer thickness. Dissolution valence as low as 0.5 can be measured for thick PS films formed at low current densities. The low n values are due to chemical dissolution of PS during the time of its growth. The thicker the PS layer the longer the sample is in the electrolyte and the larger the amount of chemical dissolution. This effect is particularly significant for the PS formed on p-Si, which has extremely fine pore structure and thus large surface area.

The formation of hydrogen does not occur in anhydrous organic solvents.[32,33,63] Due to the lack of hydrogen evolution, dissolution valence is near 4 at all current densities. Addition of water in the organic solvents reduces the dissolution valence.

3. Growth Rate of Porous Silicon

The growth rate of a PS layer can vary over a wide range depending on formation conditions. It can be as low as a few Å/s and as high as 4000 Å/s.[7,17] For p-Si the growth rate of PS at a given current density appears to increase linearly with respect to HF concentration and current density.[7,64] For n-Si the growth rate increases also with HF concentration and current density but the relations are not linear (Figure 6).[7,19,65] Also, for p-Si, it increases linearly with logarithmic dopant concentration but for n-Si the dependence of growth rate on dopand concentration is more complicated. Temperature of HF solutions generally has little

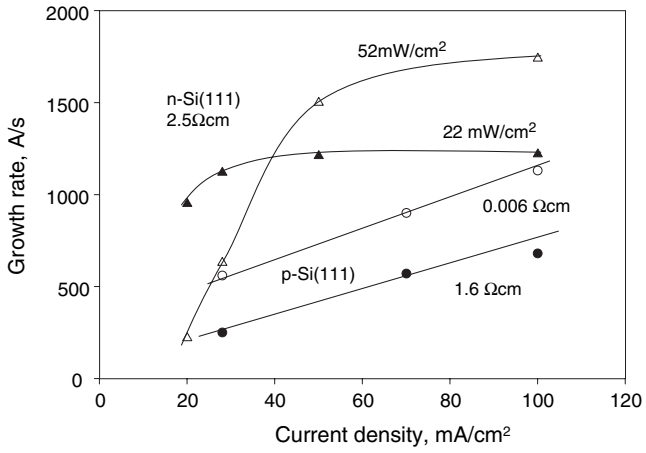


Figure 6. Effect of current density on PS formation rate, after Arita and Sunahama.[7]

influence on growth rate of PS.[2,66] PS growth rate is generally higher on (100) substrates than on (111) substrates.[67]

Figure 7 shows that PS thickness increases linearly with time up to certain thickness.[16,17] Such constant growth rate at a constant current density means that the PS formed is uniform in thickness (Effective surface area remains constant assuming reaction kinetics is the same). At a large thickness the growth may deviate from linearity due to the effect of diffusion in the electrolyte within the pores.[19,25] It has been found that for a very thick PS layer (150 μm) there is about 20% difference in HF concentration between that at the tips of pores and that in the bulk solution.[19]

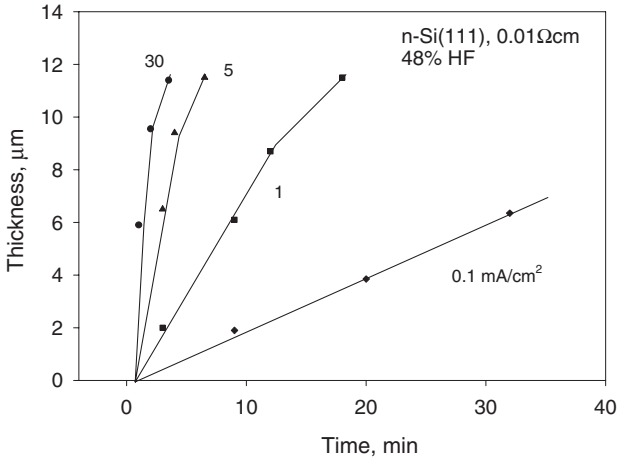


Figure 7. Kinetic dependance of the PS thickness on anodization time, after Parkhutik *et. al.*[17]

4. Chemical Dissolution During PS Formation

During PS formation at an anodic potential, the tip of pores dissolves preferentially due to easy excess to holes. The pore wall areas, which are sufficiently distant from the tips, have no holes available, dissolve chemically at a very low rate. The chemical dissolution does not depend on potential but on the time of immersion and the total surface area of the PS. However, although the dissolution rate is very low, a significant amount of material may be removed by the chemical dissolution during the formation period of PS due to the large surface area of PS. Chemical dissolution is responsible for the dissolution valence lower than 2 and the change of PS density with depth. Figure 8, as an example, shows that the amount of chemical dissolution increases with immersion time in the HF solution and with decreasing HF concentration.[68]

The data in Figure 8 can be used to estimate the chemical dissolution rate on the surface of pore walls. For a PS with a density of 50% and an average pore diameter of 3 nm, the chemical dissolution rate is estimated to be about $6x10^{-4}$ Å/s,

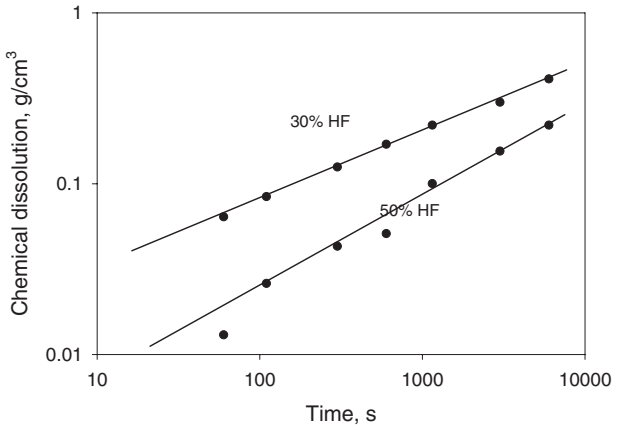


Figure 8 Relationship between amount of dissolved silicon and dipping time in HF solution for the PS formed in 30% and 50% HF, after Unno *et.al.*[68]

assuming the PS consists of straight cylindrical pores of an equal diameter. The order of magnitude is in agreement with the planar etch rate of silicon in concentrated HF solutions.

III. MORPHOLOGY OF PS

Morphology, which is determined by the distribution of materials in space, is the least quantifiable aspect of a material. It is thus very difficult to characterize morphology of PS, which has extremely rich details with respect to the range of variations in pore size, shape, orientation, branch, interconnection, and distribution. Qualitatively, the diverse morphological features of PS reported in the literature can be summarized by Figure 9 with respect to four major different aspects: pore orientation, fill of macro pores, branching, and depth variation of a PS layer.

1. Pore Diameter

Among the morphological features, the average pore diameter of a PS, as a quantifiable and easily measurable parameter, is most commonly determined. Table 1 shows some examples for pore diameter and interpore spacing reported for the PS formed under various conditions. The pores are categorized in this following text as micro pores (less than 10 nm) and macro pores (larger than 50nm). Pores with size between 10 and 50 nm are grouped into either micro or macro pores depending on the specific situation.

(*i*) *Effect of Doping*

The diameter of pores and interpore spacing may vary with doping type and concentration in a wide range, from about 1 nm to about 10 μm, about 4 orders of magnitude.[14,73,80,81] The PS formed on different substrates can be roughly grouped into four main categories according to doping concentration: 1. moderately doped p-Si (10^{15}-10^{18}); 2. heavily doped materials, p^+-Si and n^+-Si ($> 10^{19}$); 3. non- heavily doped n-Si ($< 10^{18}$); and 4. lowly doped p-Si ($< 10^{15}$). The PS formed on moderately doped p-Si has extremely small pores ranging typically from 1 to 10 nm. For heavily doped p and n types, the pores have diameters

(1) Orientation

(a) aligned to <100> and source of holes

(b) roughly aligned to source of holes

(c) partially aligned to <100> and source of holes

(d) aligned only to <100>

(2) Branching

(a) smooth pore wall

(b) branches shorter than diameter

(c) second level branches only

(d) dendritic branches

(e) main pores with second and third level branches

(f) dense, random and short branches

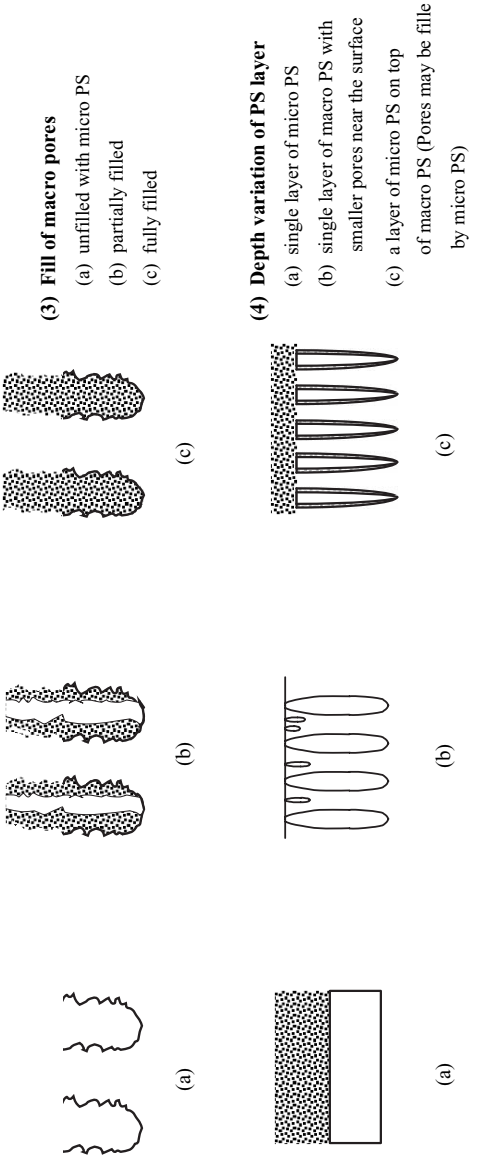


Figure 9. Morphological characteristics of porous silicon.

Table 1
Examples of Pore Diameter, Interpore Spacing and Pore Density of the PS Formed under Different Conditions (a More Extensive List is Given in Ref.[1])

Material Ω.cm	Solution % HF	Current mA/cm^2	Diameter μm	Spacing μm	Ref.
		p (100)			
0.001	20	100	0.05	0.01	59
0.006	33	50	0.005		69
0.006	24 + 47% eth	200	0.002	0.0015	70
0.01	10	10	0.0052		60
0.01	69	30	0.006	0.015	71
0.02	25 + 50% eth	30	0.05	0.02	72
1	49	30	0.003		71
13	2 M HF + MeCN	7	1–2	2–3	33
20	3	2	1.5–3, MP[c]	1–3	39
25	48	10	0.0025		12
50	10	3	2.4–4, MP	1–2	39
100	25	20	3–6, MP	0.5	73
224	2 M HF + MeCN[b]	7	1–2	0.5–2	32
		(111)			
10	50	20	0.004		74
10	50:etha.(1:3)	20	0.004		74
		(110)			
0.01	49	2	0.005		14
0.4	24 + 47% eth	50	0.035	0.01	70
		n-Si (100)			
0.002	1	4	0.05	0.05	65
0.007	15	65	0.004		60
0.01	5	0.1 V, 15 mA, illu.	1–4	0–4	27
0.03	4	4	0.05	0.3-0.8	75
0.1	49	80	0.06		76
0.1	5	6 V	0.14	0.44	24
0.1	20	50	0.009	0.03	77
0.1	20:60, HF:eth	50 mA	0.08	0.1	59
1	2.5	0.5, 10 V, illu., B	0.6	6.9	20
3	10:35, HF:eth	1.2 V, 5mA illu.	0.33	1–2	78
5	6	3 V, 10 mA illu.B[a]	2.5	5	19
10	20:60, HF:eth	50 mA	0.8	> 0.5	59
20	2.5	1.5, 10 V, illu., B	2	5.5	20

Table 1
Continuation

Material Ω.cm	Solution % HF	Current mA/cm^2	Diameter µm	Spacing µm	Ref.
		(111)			
0.01	12–48	20	0.01		79
0.008	25:50, HF:eth	50	0.01	0.03	80
2.5	50	30, 7 mW	1.5		7
2.5	50	90, 54 mW	1.7		30
		(110)			
0.82	49	80	0.01	0.1	13
3	4	4, 1.5 V	2–4	1–2	39

[a] B: backside illumination
[b] MeCN: acetonitrile
[c] MP: macropores of two-layer PS
[d] eth: ethanol,

typically ranging from 10 to 100 nm. For non-heavily doped n-Si, the pores have a wide range of possible diameters from 10 nm to 10 µm. For lowly doped p-Si, the PS can have two distinct distributions of pore diameters: large pores with a distribution of diameters in the order of µm and small pores on the order of nm. The effect of doping on pore diameter strongly depends on solution composition, potential and illumination conditions.

Figure 10 shows that for n-Si the diameter of the pores decreases with doping concentration at different current densities.[80] In contrast, the pore diameter of p-Si of moderate or high doping concentrations decreases with increasing doping concentration. In the transition region (Figure 4), macro pores can form on all types of substrates, in which the silicon surface is partially covered with an oxide film, and the morphology of PS type strongly depends on solution composition.[18,27,73,75,82,83] The macro pores formed on p-Si in the transition region have been investigated in a recent study in various solutions.[55]

(*ii*) *Effect of Potential, Current and HF Concentration*

Pore diameter generally increases with increasing potential and current density.[9,24,60,80] Figure 10 shows that the diameter of pores formed on both p-Si and n-Si increases with current

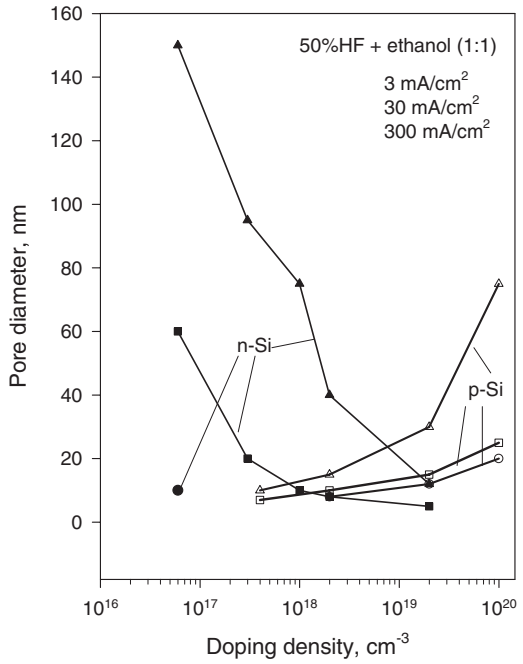


Figure 10. Pore diameter of the PS formed in 5% HF + ethanol (1:1) as a function of doping concentration and current density. After Lehmann, *et.al.*[80]

density over a wide range of doping concentrations.[60,80] Pore diameter increases, in general, with increasing potential and with decreasing HF concentration.[24] The sensitivity of pore diameter to HF concentration strongly depends on solvent.[8,27,37,39,43] A wider range of pore diameters can be obtained in organic solvents than in aqueous solutions.

(*iii*) *Primary and Branched Pores*

As illustrated in Figure 9(2), pores can be straight with smooth walls or can be branched. The branched pores can have

third level or further levels of branches. Branched and hierarchical pore structure has been found to form on all types of substrates. Branched pores are generally smaller than the primary pores.[12,14,76,84]

(*iv*) *Surface-Patterned PoreAarrays*

Straight large pores with smooth walls can be formed by backside illumination of n-Si of (100) orientation.[19,20,85] Using micro patterning for pore initiation sites on the surface regularly spaced pore arrays can be produced. Production of pore arrays of 2 μm in diameter up to 400 μm deep on 6 in wafer have been reported.[86] Well aligned pore arrays have not been obtained on (111) and (110) wafers due to the misalignment between the direction of pore growth and that of the hole source.[87]

Under a given formation condition, the pores beyond the surface region have certain size determined by the formation conditions (i.e., HF concentration, current density, potential and doping concentration). According to Lehmann et al.,[87] the smallest possible pore diameter for a regular pore array formed with surface patterning and backside illumination is about 0.3μm below which branching at pore bottom occurs; the largest pores are found to be about 20 μm, above which formation of straight and smooth pores becomes a problem due to hydrogen bubble formation.

(*v*) *Variation From Surface To Bulk*

The pores at the surface are smaller than those in the bulk of PS as, for example, shown in Figure 11.[8,16,24] Such an increase in pore diameter from the surface to bulk is due to the transition from pore initiation to steady growth. Also, two-layer PS, a micro PS layer on top of a macro PS can form for on illuminated n-Si or on lowly doped p-Si.[19,26,78] For the micro PS layer formed on front illuminated n-Si, pore diameter is less than 2 nm and thickness of PS changes with illumination intensity and the amount of charge passed. Also, the diameter of macro pores on front illuminated n-Si changes with the amount of charge Passed.[20] Pore size and depth variation of PS on n-Si are very different for front and back illuminated n-Si samples.

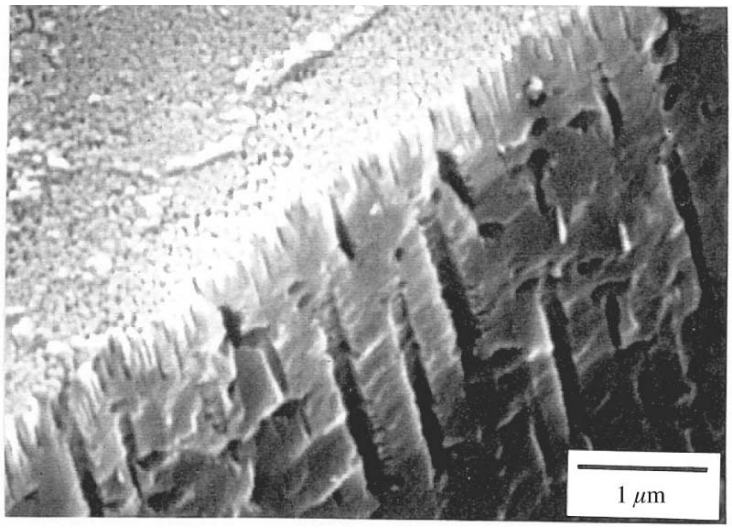


Figure 11. SEM photograph showing smaller pore diameter and larger pore density near the surface than in the bulk in 5% HF at 6 V.[24]

For very deep pores the diameter may increase with a decreased growth rate due to the effect of diffusion process inside pores. The depth at which this occurs depends on current density and HF concentration.[87] Low HF concentration, low temperature and low growth rate favour formation of deep uniform pores.

(*vi*) *Interpore Spacing*

The variation of interpore spacing or wall thickness is more complex than pore diameter. There is not much information on the relationship between wall thickness and formation conditions. In general, the walls are on the same order as or smaller than pore diameter.[24] In particular, interpore spacing depends on potential; it increases with potential at small currents but at certain current it starts to decrease with increasing potential. When interpore spacing is reduced to zero, which occurs in the transition region, pores no longer form, and instead, shallow pits form.

(*vii*) *Distribution of Pore Diameter*

Pore diameters of the PS layer formed under a given set of conditions have a distinct distribution. Normal, log-normal, bimodal, fractal, and non-uniform distributions have been found for PS formed under different conditions.[26,29,60,88,89,90,91]

For the PS formed on heavily doped silicon, pore diameter has a narrower distribution at a lower current at a given HF concentration and the distribution is narrower at a lower HF concentration at a given current desnity. Bimodal distribution of pore diameter are generally associated with two layer PS on lowly doped p-Si and illuminated n-Si. Pores with multiple distributions have been observed for the PS that has a surface micro pore layer and smaller branched pores in addition to the main pores.[88,90] Distribution of pore diameters for highly branched PS has been found to be fractal-like.[14,76] Illumination during formation of PS on p-Si can also to affect the distribution of pore diameter; it increases the amount of the smaller nanoscrystals, while reduces the amount of larger crystals.[92] For the PS formed under an illuminated substrate, the relative amount of small crystals tends to increase with reduction of light wavelength.[93]

2. Pore Orientation and Shape

The growth of individual pores, or the dissolution at pore tips, is anisotropic, depending on the orientation of the substrate and the direction of the carrier source. The relative effect of substrate orientation and direction of the carrier source on the orientation of pores is determined by the specific electrochemical reactions at the pore bottoms.

The orientation of primary pores is in general in the <100> direction for all the PS formed on all types of (100) substrates.[7,14,84] For the PS with dendritic structure, as shown in Figure 12, pores propagate along (100) direction even on the (110) and (111) substrates.[20,84,94] The branched pores of non-dendritic types formed on (100) substrate may not be strictly perpendicular to the primary pores but deviate to various extents from the <100> direction toward to the source of holes.[12,14,95] Ronnebeck et al.[95] found that the macro pores formed on

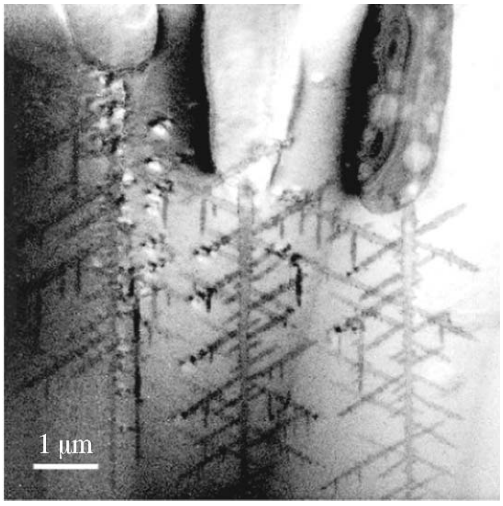


Figure 12. Dendritic PS morphology formed on n (100), 12 Ω cm in 4% HF at 4 mA/cm^2. After Jäger, *et. al.*[84]

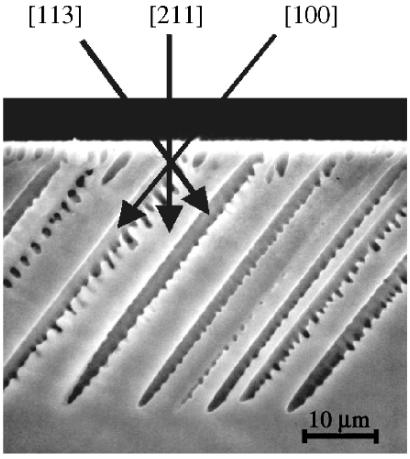


Figure 13. Preferential growth of pore on 35^{o}C from (100) n-Si with back illumination. After Ronnebeck *et. al.*[95]

misoriented n(100) substrates (from 15^0 to 35^0) as well as on n(111) are oriented along <100> direction when they are formed in the dark at high anodic potentials. However, when the samples are illuminated from the back, the pores formed on the (111) sample and the side pores formed on the 35^0 misoriented (100) sample are oriented along <113> as shown in Figure 13.

For pores of extremely small diameters, in the order of a few nm, the direction of individual pores is totally random. On the other hand, large pores tend to have less anisotropic effect and grow more dominantly in the direction of carrier supply, that is, perpendicular to the surface. The macro pores formed on p-Si generally have smooth walls and an orientation toward the source of holes that is perpendicular to the surface, even on (110) and (111) samples.[34,39]

Depending on substrate orientation and formation condition, individual pores may have different shapes. The shape of the pores formed on (100) substrate is a square bounded by {011} planes with corners pointing to the <100> directions.[14,77] The shape of individual pores formed on n-Si tends to change from circular to square to star-like and to dendrite-like with increasing potential.[20] Low formation voltage tends to favour circular shape while high voltage favours star-like shape. Near perfect square shape of pores can be obtained for the PS formed on n-Si under certain conditions.

The bottom of individual pores is always curved, varying from a shallowly curved semicircle to an elongated conical depending on the formation conditions. As will be discussed later the curvature of pore bottom plays a critical role in the reaction kinetics required for formation of PS and its morphology.

3. Pore Branching

Individual pores, depending on formation conditions, may propagate straight in the preferred direction with very little branching or with formation of numerous side or branched pores. In general, the conditions that favour the formation of small pores also favours branching.

The degree of branching and interpore connection depends strongly on doping concentration. The most highly connected PS

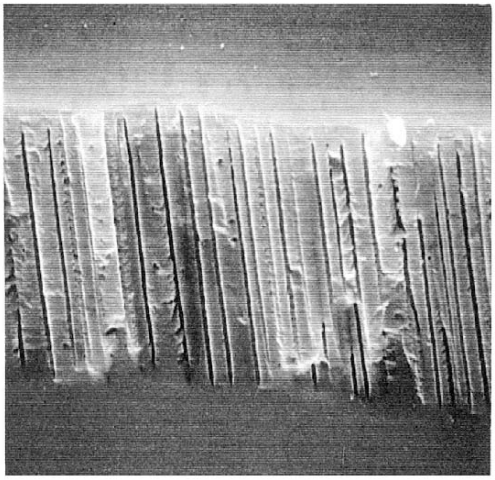


Figure 14. Straight pores with smooth wall formed on n-Si in the dark.[1]

is found in the PS of lowly doped p-Si and the micro PS on illuminated n-Si.[14] On the other hand, well separated and straight pores are generally found on moderately or lowly doped n-Si. Perfectly straight pores with smooth walls can be formed on n-Si in the dark, as for example, shown in Figure 14.[65] Also, on n-Si smooth and straight pores without branching can be obtained under back illumination using surface patterned substrate. For heavily doped materials, pores are generally branched.[12,71] For non-heavily doped p-Si, pores are highly branched resulting in a random network of pores. The macro pores formed on lowly doped p-Si generally have no side pores longer than the pore diameter.[32,34,39]

Like main pores, branched pores are highly directional, propagating preferentially in the {100} planes and the long <100> directions.[14,20,77] Such directional branching can produce regularly spaced three dimensional structures as shown in Figure 15.[14,65] The tendency to branch is stronger on (110) and (111) substrates than on (100) because for substrates of non-(100) orientations, <100> direction does not coincide with the direction of the hole source. Due to the large tendency to branch, it is difficult to

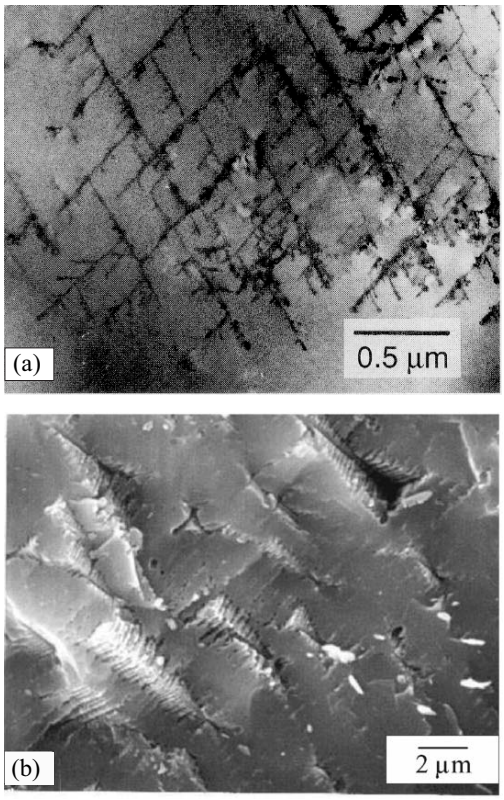


Figure 15. Three dimensioned structures generated by the formation of dendritic pores on an n(110) substrate.[1,14]

produce straight perpendicular pores with smooth walls on (110) and (111) samples, even on surface patterned samples.[87]

4. Interface Between PS and Silicon

The growth of a PS layer (not individual pores) is always perpendicular to the surface of the substrate with back electrical contact. The growth front is planar, independent of the orientation of the silicon samples. The interface between PS layer and silicon

substrate is essentially flat with high and low points deferring within a few microns.

5. Depth Variation

The morphology of PS generally varies in the depth direction from the surface to the bulk. There are two types of depth variations: 1) the change of pore diameter is gradual from the surface to the bulk where pore diameter is constant; and 2) the change of pore diameter is abrupt from a surface layer to the bulk layer with a difference in diameter as large as three orders of magnitude. The surface layer for the first type is a transition layer formed due to the transition from initiation of pores to the steady state growth. The second type is two-layer PS with a micro PS on top of a macro PS. While two-layer PS forms only under certain conditions, the transition PS layer of varying thickness exists on the surface of all PS layers.

(*i*) *Transition Layer*

Transition layer is found to exist for all types of silicon.[7,16,20,24,25,80] The pores in the transition layer are generally much smaller than those in the bulk. There is not a clearly definable boundary that separates the surface layer and the bulk. The thickness of the transition layer is related to the size of pores; the smaller the pores the thinner the surface transition layer. For n-Si, the transition layer can be clearly seen as for example shown in Figures 11 and 16.[24] On the other hand, for p-Si this surface layer is very thin (near zero) for some PS with extremely small pores. Such thin layer may not be observed because it may be removed due to chemical dissolution during its exposure in solution.

The morphology of the transition layer, unlike the bulk morphology, depends sensitively on surface conditions, particularly surface roughness such as scratches.[14,94] For n-Si, which usually requires a large potential to generate current in the dark, formation of PS can occur at much lower potentials if the surface is roughened mechanically.

A certain amount of uniform dissolution may occur prior to and during the initiation of pores. Such an etched layer prior to the

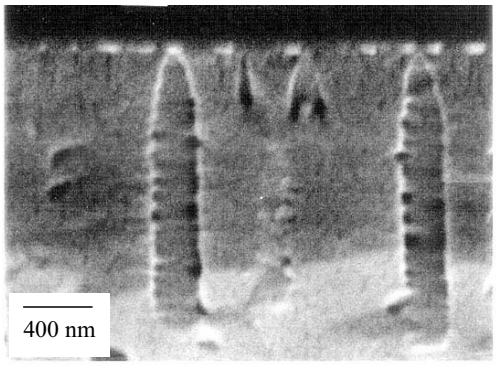


Figure 16. Variation of diameter of pores from the surface to the bulk. After Zhang.[24]

initiation of pores is involved in all types of PS since the etching, which causes roughening of the surface, is required for pore initiation.[84,95] The thickness of etched layer is generally in the order of pore diameter.

(*ii*) *Two-Layer PS*

Two-layer PS, with a micro PS layer on top of a macro PS layer, forms only under certain conditions. For n-Si, formation of two-layer PS is associated with front illumination, although it can also be formed with back illumination.[19,27,34,37] For p-Si, two-layer PS are found to form on lowly doped substrates. For moderately or highly doped p-Si or for n-Si in the dark, formation of two-layer PS has not been observed.

Formation of two-layer PS on front illuminated n-Si depends on the wavelength of light; formation of large and deep pores does not occur with a white light at wavelength < 800 nm, but occurs at wavelength > 867 nm.[20] This phenomenon is due to the wavelength dependent absorption depth; carriers which are generated deep in the bulk promote pore growth at the tips of macro pores whereas those generated near surface result in the formation and dissolution of micro PS and lateral growth of the macro pores.

There is a correlation between the occurrence of two-layer PS and the saturation photo current value.[78] Only a single micro PS layer forms at a photo current density below the photo saturation value while two-layer PS forms at current densities above the saturation current as shown in Figure 17. Also, macro PS layer forms only after a certain amount of charge determined by the amount of etch required for the initiation of macro pores.[77]

There is also an etched layer of Si on the surface under illumination as illustrated in Figure 18. This etched layer is mainly due to photo-induced corrosion. As a result of the photo induced dissolution the top surface of PS layer recedes with time. The rate of dissolution depends on doping, HF concentration, current density and illumination intensity.

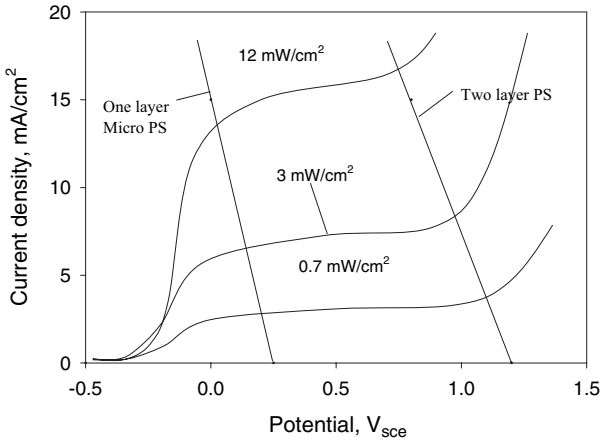


Figure 17. Voltammograms of n-Si electrodes in aqueous solution containing 10 wt.% HF and 35 wt.% C_2H_5OH at 5 mV s^{-1} under different light intensities. The illumination intensity was adjusted by changing the distance between the light source and the n-Si electrode. After Osaka *et.al,*[78]

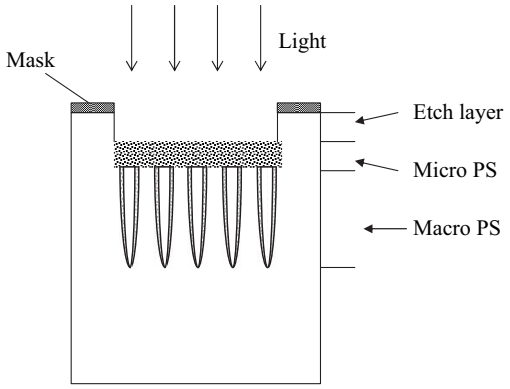


Figure 18. Three layers are resulted after formation of PS on illuminated n-Si sample: an etch layer, a micro PS layer and a macro PS layer. The walls of the macro pores may also be fully or partially covered with micro PS.

The micro PS on top of the macro PS formed on n-Si under illumination can have a very complex structure. It has been reported that the PS formed on n-Si has an irregular and square-shaped framework of pores with interconnected columns of 100-200 nm for lowly doped materials and 100-400 nm for highly doped materials.[29] Bounded to this rigid square-shaped framework and contained within is a finer structure with wire-like structures of sizes varying from 3 to 12 nm. Fractal morphology has been suggested to associated with this type of micro PS. The macro pores formed on front illuminated n-Si, once initiated, can increase in diameter with time without reaching a constant value.[20,27] The diameter of macro pores also increases with increasing HF concentration.

The PS formed on p-Si generally consists of micro pores, but when the resistivity is above a certain value macro pores can form underneath a layer of micro pores.[34,39,73] The resistivity, at which this occurs, depends on type of solvent, HF concentration and

current density. In aqueous HF solutions macro pores are found to form on substrates of resistivity higher than 5 Ωcm.[39] On the other hand, macro pores has been found to form on substrate with a resistivity of 1 Ωcm in 2M HF + dimethylformamide electrolyte.[34] Presence of water in organic solvents tends to reduce the thickness of the micro PS layer.

(*iii*) *Fill of Pores*

The macro pores of two-layer PS may be completely filled or only partially filled with micro pores depending on the solution composition.[39,63] When other conditions are the same, macro pores formed in organic solvents tend to be more filled than those formed in aqueous solutions. Addition of water to organic solutions reduces the extent of filling. Among organic solutions, according to Jager et al.,[54] the degree of filling of the macro pores formed on p-Si depends on the oxidizing nature of the solutions: macro pores are filled with micro PS in non-oxidizing electrolyte such as acetonitrile (MeCN), while they are not filled for oxidizing electrolytes such as dimethylformamide (DMF).

For the two-layer PS formed on front illuminated n-Si, the degree of filling of the macro pores depends on light intensity, on the amount of charge passed, as well as on the magnitude of the current relative to the saturation photo current.[26,27,29,78] The thickness of micro PS increases with the amount of charge passed and it decreases with increasing light intensity.[27] The macro pores formed on lowly doped n-Si under back illumination can also be filled by micro PS and the degree of filling decreases with increasing potential.[19]

6. Summary

The morphology of PS has extremely rich details determined by the numerous factors involved in the anodization. Generally, p-Si and n-Si have distinct differences in the correlation between PS morphology and formation conditions. Among all formation conditions doping concentration appears to show the most clear functional effect on morphology. A summery of morphology features of PS is provided in Table 2.

Table 2
Morphological Characteristics of Porous Silicon[a]

Size

1. A PS layer may have one or two distinct distributions of pores with sizes varying from 1nm–10μm. Those less than 10 nm are micro pores and those larger than 50 nm are macro pores.
   In the dark
   moderately doped p-Si (10^{15}–10^{18}): one distribution, 1–10 nm.
   heavily doped p-Si and n-Si (> 10^{19}): one distribution,10–100 nm.
   non-heavily doped n-Si (< 10^{18}): one distribution,10 nm–10μm.
   lowly doped p-Si (< 10^{15}): one or two distributions, 1–10 nm and > 1μm.
   Illumination
   Little effect on p-Si.
   For n-Si, one or two distributions 1–10 nm and 50 nm–10 μm.
2. For p-Si pore diameter increases with doping concentration; for n-Si, it decreases with doping concentration.
3. It increases with potential or current density but decreases with HF concentration.
4. Branched pores are smaller than primary ones.
5. It is constant in the bulk but is smaller near the surface.
6. Inter-pore spacing can be smaller than pore diameter but is not larger than trice the diameter. It varies with the factors similarly as pore diameter.

Orientation

1. Pores grow preferentially along <100> directions and toward the source of holes.
2. Well aligned macro pores, > 50 nm are perpendicular to the surface for (100) substrate but may have a less than 90^0 angle to the surface on other substrates.
3. Pores can be fully aligned with <100> directions or/and with source of holes. Pores with smooth wall tend to be aligned with source of holes, while dendritic pores are aligned with <100> directions.
4. Very small pores, < 10 nm, do not show clear orientations.

Branching

1. PS has discrete pores with no branches, with short branches, or with dendritic branches, or have no primary but densely and randomly branched pores.
2. Tendency to branch increases with decreasing pore diameter.
3. Direction of branching has the same tendencies as pore orientation.

Fill of pores

1. In PS of two distributions, macro pores may be partially or fully filled by micro pores.
2. It tends to occur in organic solvents; addition of water reduces the extent of filling.
3. Extent of filling decreases with increasing light intensity or increasing potential.

Table 2. Continuation

Depth variation
1. Two types: transitional layer, which has no clear boundary with bulk and two layer type with clear boundary between micro and macro PS layers.
2. Transitional layer is associated with the initiation of pores with smaller pores at the surface.
3. Two layered PS occurs only on illuminated n-Si and lowly doped p-Si.
4. Pore diameter and distribution is uniform in the bulk of PS.
Others
1. The bottom of pores is curved with the smallest curvature at the tips.
2. The shape of pores can be square, dendritic, circular, and star-like.
3. Interface of PS/Si is flat and is perpendicular to the source of holes.
4. PS with macro pores has the same composition and crystalline structure as the substrate; PS with micro pores can be amorphous and highly hydrogenated and/or oxidized.
5. Pore wall is rough at atomic scale for all pores.

[a] Based on information summarized in Ref. [1].

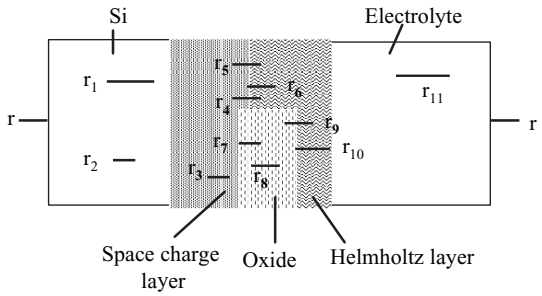
IV. ANODIC REACTION KINETICS

1. Potential Distribution and Rate Limiting Process

There are five possible physical phases in the current path in which the current conduction mechanisms are different as illustrated in Figure 19. They are substrate, space charge layer, Helmholtz layer, surface oxide film, and electrolyte. The overall change in the applied potential due to a change of current density in the current path is the sum of the potential drops in these phases:

$$\Delta V_{app} = \Delta V_{Si} + \Delta V_s + \Delta V_H + \Delta V_{ox} + \Delta V_{el} \quad (1)$$

The distribution of the applied potential in these different phases depends on silicon material as well as potential range. Table 3 is a summary for the potential distribution of different materials in the three potential regions (Figure 2). While the distribution of the potential below the passivation potential V_p (potential at current peak J_1 in Figure 2) depends on doping type and concentration,



r_1 & r_2 - majority & minority carrier transport,
r_3 - transport of holes to the surface
r_4 - charge transfer across the Helmhotz layer,
r_5 - electron injection
r_6 - chemical dissolution,
r_7 - oxide formation,
r_8 - ionic transport in oxide,
r_9 - injection of oxidants,
r_{10} - dissolution of oxide,
r_{11} - mass transport in electrolyte

Figure 19. Schematic illustration of the processes involving the transport of charge and species in the different phases in the interface region.

Table 3
Distribution of the Applied Potential in the Electric Layers at the Silicon/Electrolyte Interface in HF Solutions

Material	OCP to V (i ~ 0)[a]	Exponential region	$V > V_p$
Heavily doped p, n-Si	Helmholtz layer	Helmholtz layer	Oxide film
Non-heavily doped n-Si	Space charge layer	Space charge layer	Oxide film
Non-heavily doped p-Si	Space charge layer	Helmholtz layer & space charge layer	Oxide film

[a]OCP, V (I ~ 0), V_p: defined in Figure 2.

the applied potential is principally dropped in the oxide film at a potential higher than V_p for all materials.

For n-Si the applied potential is predominantly dropped across the space charge layer in the semiconductor. For p-Si and heavily doped n-Si, distribution of potential is partitioned between the Helmholtz layer and/or the space charge layer.[18,48-51] For heavily doped materials the surface is degenerated and the material behaves like a metal electrode, meaning that the charge transfer reaction in the Helmholtz double layer is the rate determining step and potential drops mostly in the Helmholtz layer. For non-heavily doped p-Si, the potential is mostly dropped within the space charge layer before the onset of current. At potentials higher than that at which current becomes measurable, the change in potential may be dropped also in the Helmholtz layer in addition to the space charge layer. Also, for very lightly doped material, potential drop in the substrate can also be significant.[52]

Because of the different potential distributions for different sets of conditions the apparent value of Tafel slope, about 60 mV, may have contributions from the various processes. The exact value may vary due to several factors which have different effects on the current-potential relationship: 1) relative potential drops in the space charge layer and the Helmholtz layer; 2) increase in surface area during the course of anodization due to formation of PS; 3) change of the dissolution valence with potential; 4) electron injection into the conduction band; and 5) potential drops in the bulk semiconductor and electrolyte.

Each of the possible processes in anodic dissolution of Si, as illustrated in Figure 19, can be the rate limiting process under certain conditions. For example, the anodic reaction processes on n-Si in the dark is limited by the minority hole transport in the bulk of silicon, that is r_2. For p-Si and illuminated n-Si in fluoride solutions at potentials negative of the first current peak, J_1, the reaction rate is determined by the charge transfer process across the electrode/electrolyte interface, that is, r_4 and r_5. At potentials positive of J_1, i.e., the electropolishing region, the rate determining step is the dissolution of the anodic oxide film, that is r_{10}. The dissolution of the oxide film formed at low fluoride concentrations is mainly kinetically controlled, that is r_{10}, while at high fluoride concentrations the process is mainly diffusion

Table 4
Rate Limiting Steps Involved in the Anodic Reactions[d]

	r[a]	p	n	p^+,n^+
		Dark		
OCP	~ 0	r_3	r_3	r_3
OCP < V < V_p	large	r_4	r_2 or r_5[b]	r_4
V ≥ V_p	large	r_{10}, r_{11}	r_{10}, r_{11}	r_{10}, r_{11}
		Illuminated		
OCP	large	r_4	r_3	r_4
OCP < V < V_p	large	r_4	r_3 or r_4	r_4 or r_5[c]
V ≥ V_p	large	r_{10}, r_{11}	r_{10}, r_{11}	r_{10}, r_{11}

[a] oxidation rate, qualitative comparison to the dark limiting current density
[b] it is r_2 bellow the breakdown potential but it is r_4 above the breakdown potential;
[c] it is r_3 when the current equals the saturation photo current, is r_5 when the current is larger than the saturation photo current and is r_4 when the current is less than the saturation photo current;
[d] in aqueous solutions containing no redox couples other than the ones associated with HF and H_2O; presence of redox couples can affect the rate limiting process.

controlled, r_{11}. There is a critical concentration, depending on pH, convection and potential, at which the contributions by kinetic and diffusion processes are equal. Table 4 is a summary for the rate limiting steps under various experimental conditions.

2. Reaction Paths

The surface of silicon in HF solution is terminated by hydrogen with or without active dissolution.[53,54] Thus, dissolution of a surface silicon atom involves first the replacement of a surface hydrogen atom with F^- in HF solutions. As shown in Figure 20, replacement of H by F^- requires a hole which results in a neutralized Si-F bonding. The valence state of the adsorbed hydrogen before and after the replacement, being zero, is not changed. On the other hand, hydrogen adsorption onto a silicon atom is a reduction process since the valence of the hydrogen atom is changed from +1 to zero. It occurs when the Si-SiF bond

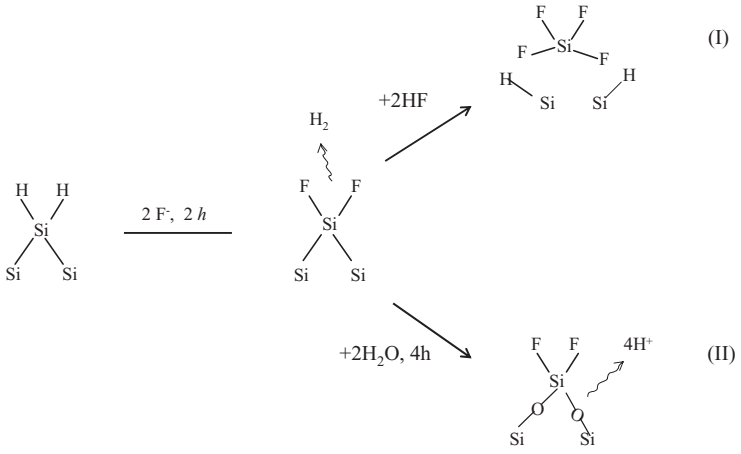


Figure 20. Two different reaction paths for a silicon electrode in HF solution

is broken by reacting with HF (path I). Transfer of one electron from the Si-SiF bond to the hydrogen from a HF occurs to form a Si-H bond.

An important feature of step (2) is that no carriers from the solid are involved although it is an oxidation process for the silicon atom involved (from valence 2 to 4). and it therefore is a chemical reaction in nature. This is the key reaction step responsible for the chemical character of the dissolution process in that in the process of becoming adsorbed a hydrogen ion is reduced by receiving an electron from the Si-SiF bond. Alternatively in the second step, the backbond of Si-SiF can be broken by reacting with H_2O, resulting in Si-O-Si bonds, which are not stable in HF and the dissolution of which results in an indirect dissolution path (path II). Thus, while reaction path (I) results in the direct dissolution of silicon, path (II) results in indirect dissolution. Note also that reaction path (I) results in a dissolution valence of 2 while that of reaction path (II) is 4. As the coverage of the surface by Si-O-Si bonds increases with increasing potential, the surface becomes increasingly less active and becomes passivated when these bonds fully cover the surface

(i.e., formation of a continuous oxide film). Further reaction has to proceed via the breaking of Si-O-Si bonds, the rate of which increases rapidly with increasing HF concentration.

Thus, the occurrence of different regions during anodization of silicon in HF solutions is associated with the two competing reaction paths: direct dissolution of silicon and indirect dissolution through formation and dissolution of silicon oxide. Formation of PS, which is associated with semiconductor property of the surface as will be detailed in a following Section, is only possible when the surface is not fully covered by an oxide film such that direct electrochemical dissolution of silicon is possible. At low potentials, direct dissolution of silicon dominates and PS is readily formed. On the other hand, at potentials higher than the current peak, the entire surface is covered with an oxide film. The dissolution is determined by the properties of the oxide film and no longer by the semiconducting properties of silicon. PS does form under such condition. The formation of an oxide film and its surface coverage determines the occurrence of these three regions shown in Figure 4. (Electropolishing region is absent in anhydrous organic solutions due to the lack of water that is required for the formation of oxide film).

The electrochemical reactions and processes involving the anodic dissolution of silicon in HF solutions have been extensively studied in the past. Table 5 provides a summary for the characteristics of the anodic processes that are relevant to the formation of PS (details are documented in Ref.[1].)

3. Effect of Geometric Elements

The overall formation mechanism of PS must involve the fundamental electrochemical reactions in three essential aspects: 1. nature of reactions, reactants, products, intermediates, number of steps, and their sequences, 2. nature and rate of charge transport in the different phases at silicon/electrolyte interface, 3. spatial and temporal distributions of reactions and the cause of such distributions. The first and second aspects, which governs the properties of a uniform and flat surface and do not involve geometric factors, have been characterized in previous Sections and the major characteristics are summarized in Table 5. This Section deals with the third aspect, that is, spatial and temporal

Table 5
Characteristics of Anodic Processes Relevant to the Formation of PS

1	In aqueous solutions silicon reacts spontaneously with water to form an oxide film which passivates the surface.
2	Presence of HF results in the dissolution of silicon oxide and activates the surface. Fluoride species such as HF and F^- also react directly with the bare silicon surface.
3	In aqueous HF solution, the dissolution of silicon atoms has two principal competing paths, one via the reaction with HF and the other with H_2O.
4	Silicon, having 4 valence electrons, dissolves in multiple steps, each of which may occur at different energy levels and may proceed via the conduction band as well as the valence band depending on the condition of Si/electrolyte interface.
5	The surface of silicon during anodic dissolution is dynamically terminated by hydrogen; the dissolution of silicon atoms proceeds by first forming a Si-H bond.
6	The replacement of the hydrogen by fluoride polarizes and weakens the Si-Si backbond and facilitates the subsequent attack on this bond by HF or H_2O.
7	The quantum efficiency of photo electrochemical reactions may vary from 2 to 4; the effective dissolution valence may vary from 2 to 4; and the efficiency of hydrogen evolution may vary from zero to near 1 depending on light intensity and potential.
8	Anodic oxide films formed under different kinetic conditions vary in structure, composition and property (e.g., etch rate) and they change with time during the anodization.
9	The applied anodic potential may mostly or partially drop in the space charge layer or in the Helmholtz layer depending on doping type and concentration as well as on the potential range.
10	The rate of removal of surface silicon atoms by the electrochemical reactions is orientation dependant, lowest on (111) and higher on other orientations.

distributions of reactions in relation to the geometrically dependent elements and events in the system. It is this aspect that determines the specific morphology of PS formed under a given condition. Table 6 is a summary of the effect of geometric factors on the kinetics of spatially distributed reactions responsible for formation of pores and the diverse morphological features of PS.

Table 6
Effect of Geometric Factors on Reaction Kinetics

Geometric factors	Effect on reaction kinetics
Surface curvature	Distribution of carriers and active atoms on surface and thus type and rate of reactions
Surface lattice structure	Density of active surface atoms and reactivity of the surface determined by the crystalline orientation of silicon/electrolyte interface
Pore size	Geometric stability of surface against random perturbations
PS thickness	Mass transport and chemical dissolution of PS
Separation of pores	Distribution of potential in the current paths of neighbouring pores
Location of carrier source	Spatial distribution of carriers at Si/electrolyte interface

(*i*) *Surface Curvature*

A particular important property of silicon electrodes (semiconductors in general) is the sensitivity of the rate of electrochemical reactions to the radius of curvature of the surface. Since an electric field is present in the space charge layer near the surface of a semiconductor, the vector of the field varies with the radius of surface curvature. The surface concentration of charge carriers and the rate of carrier supply, which are determined by the field vector, are thus affected by surface curvature. The situation is different on a metal surface. There exists no such a field inside the metal near the surface and all sites on a metal surface, whether it is curved not, is identical in this aspect.

Physically, the sensitivity of reactions to surface curvature can be associated with the space change layer or the resistance of the substrate. For moderately or highly doped materials, this sensitivity is only associated with the space change layer because the ohmic potential drop in the semiconductor substrate is very small. However, for lowly doped material a significant amount of potential can drop in the semiconductor to cause the current flow inside semiconductor to be also sensitive to the curvature of the surface.

(*ii*) *Effect of Width of Space Charge Layer*

The bottom of all pores is curved, which affects the field at the surface of a semiconductor. For an interface with a spherical shape as illustrated in Figure 21, the field in the semiconductor can be calculated by solving the Poisson's equation.[8]

$$\xi(x + r_0) = \frac{qN_b}{3\varepsilon}\left[-(x + r_0) + \frac{(r_0 + x_d)^3}{(x + r_0)^2}\right] \quad (2)$$

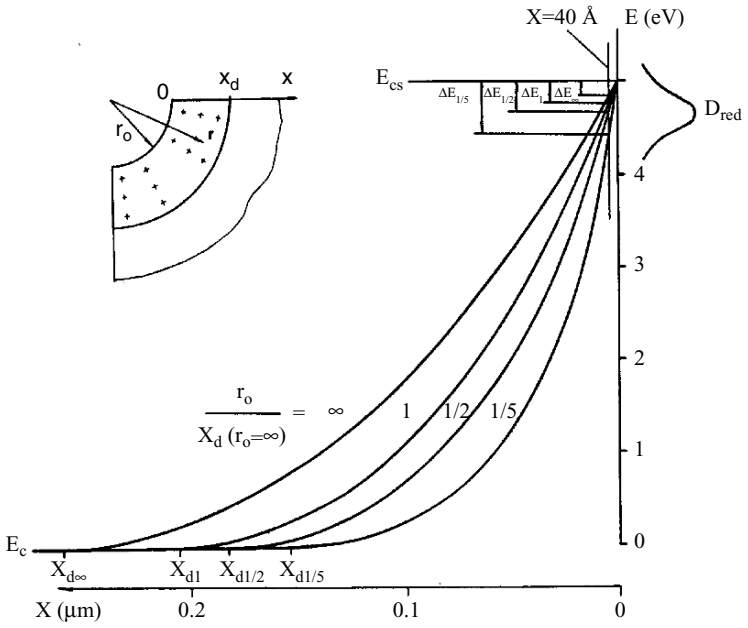


Figure 21. The energy band diagram (only the conduction band is shown) calculated for the silicon/electrolyte interface with a potential drop of 5 V and different radii of curvature. E_c is the conduction bandedge in the bulk and E_{cs} is the conduction bandedge at the surface. ΔE_4, ΔE_1, $\Delta E_{1/2}$, and $\Delta E_{1/5}$ are the possible tunneling energy ranges for different radii of curvature. The distribution of occupied states at the interface, D_{red}, is also schematically indicated. After Zhang.[24]

strength of a curved surface increases considerably when the radius of curvature is close to or smaller than the width of the space charge layer of a flat surface. Figure 21 shows the effect of in which $r = x + r_0$, N_B is the ionized donor density, q the electronic charge, r_0 the radius of curvature of the interface and x_d the space charge layer width. According to Eq. (2) the field radius of curvature on the energy band diagram with a potential drop of 5 V across the space charge layer of an n-Si under an anodic bias. Clearly the potential drops more sharply for a smaller radius of curvature and the width of space charge layer is considerably reduced. As a result, the field in the space charge layer is greatly increased, for example, the field is about 4 times higher for a curved surface with a radius of curvature of $1/5\ x_d(r_0 = 4)$ than that of the flat surface. For n-Si under anodic potentials the current conduction is by electron tunnelling from the surface into the conduction band. Note that the steep band bending due to decreasing radius of curvature increases the number of energy levels for electron tunnelling. In the case shown in Figure 21 the available energy range for tunnelling at a tunnelling distance of 40 Å is about 180 mV for a flat surface while it is about 470 mV for a curved surface with a radius of curvature of $1/5\ x_d(r_0 = 4)$.

Similar analysis can be made for other types of materials. Thus, as a generalization, the curvature of a surface causes field intensification, which results in a higher current than that on a flat surface. Although the detailed current flow mechanism can be different for different types of materials under different potentials and illumination conditions, the effect of surface curvature on the field intensification at local areas is the same. The important point is that the order of magnitude for the radius of curvature that can cause a significant effect on field intensification is different for the substrates of different widths of the space charge layer. This is a principle factor that determines the dimensions of the pores.

4. Potential Drop in the Substrate

For a moderately or highly doped material the potential drop due to ohmic resistance in the substrate is very small. For example, for a substrate with a resistivity of 0.1 Ωcm ($\sim 10^{17}/cm^3$) the potential drop in the substrate of 0.1 mm thick at a current density of 10

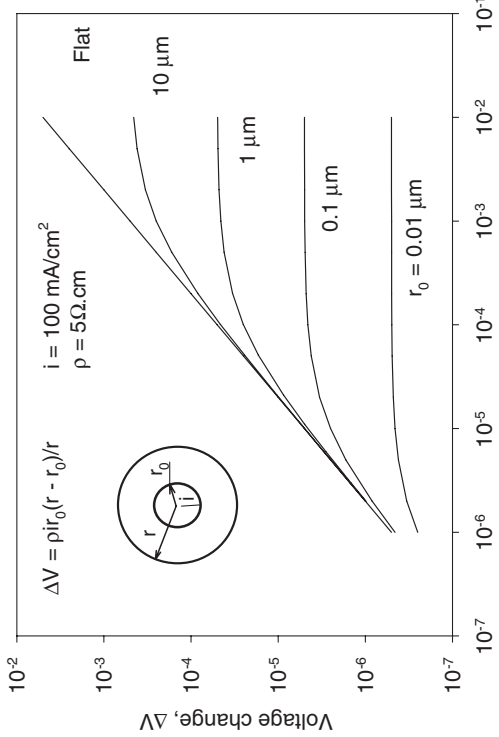


Figure 22. The change of voltage due to resistance of the material in the substrate of a hollow sphere as a function of radius of curvature of the inner sphere.

mA/cm^2 is 0.01 mV, which is insignificant compared to the potential of millivolt range that is required to significantly affect the rate of charge transfer process at the interface. However, when the resistivity is on the order of 10 Ωcm ($\sim 10^{15}/cm^3$) or larger the potential drop inside the substrate is in the mV level which starts to affect the distribution of the change of potential in the current path.

For a solid between two spherical surfaces shown in Figure 22, the resistance can be described by

$$R = \rho(r-r_0)/(4\pi r_0 r) \tag{3}$$

where ρ is the resistivity of the material, r_0 is the diameter of the inner sphere and r the outer sphere. The potential drop in the substrate, ΔV_s, for a current $I = i4\pi r_0^2$, where i is the current density on the inner surface, flowing though the solid is then described by

$$\Delta V_s = IR = \rho i r_0(r-r_0)/r \tag{4}$$

Figure 22 shows that the potential drop in the material with a curved surface increases non-linearly with increasing distance from the inner sphere and most of the potential drop occurs within a distance of a few times the diameter of the inner sphere. Also, for the same thickness of the solid the total potential drop increases with increasing radius of curvature of the inner sphere, reaching the maximum at an infinitely large radius of curvature, that is, a flat surface. The value of the potential drop is in the order of mV for a inner diameter of 10 μm under the conditions shown in Figure 22. It will be larger if the current density or the resistivity is higher. Also, the results shown in Figure 22 are for a spherical surface. In the case of curved pore bottoms, which are roughly semi-spherical, the potential drop from individual pore bottom into the solid must be significantly larger than that of an isolated sphere because the current flow from one pore overlaps with those from the neighbouring pores.

The results of the above analysis suggest that the formation of macro PS on lowly doped materials can be associated with a non-linear potential distribution in the solid of a curved surface due to the high resistivity of the solid. The formation of two-layer PS on

p-Si indicates that there are two different physical layers in which the potential-current relations are sensitive to the radius of curvature. The space charge layer of p-Si is thin under an anodic potential and is associated with the formation of the micro PS. The non-linear resistive effect of the substrate is responsible for that of macro PS. Also, to have the same change of potential drop in the substrate, ΔV_s, at a given current density, the radius of curvature must be smaller for a material of larger resistivity according to Eq. (3). This means that the pore diameter decreases with increasing resistivity, which agrees with experimental results.[39]

As a further deduction, the effect of high substrate resistivity should also occur in the case of lowly doped n-Si. However, the width of the space charge layer in such a substrate under an anodic potential is on the same order of magnitude as the dimension of the resistive layer. Since macro PS forms on n-Si substrate at an anodic potential in the dark, the effect of these two causes on the diameter of the pores formed on n-Si in the dark are not distinguishable under normal conditions. If the conditions can be controlled in such that the pores formed due to the effect of the space charge layer and that due to the effect of the resistive layer have a size difference by at least one order of magnitude it would be possible to obtain the PS with two distinct distributions of pore diameters on high resistivity n-Si. Such a condition might exist at a low potential (small space charge layer thickness) on a back illuminated n-Si substrate. It might even be possible to obtain the PS with three distinct distributions of pore diameters on a front illuminated n-Si sample.

5. Surface Lattice Structure

The surface of a single crystal can have different reactivities depending on the orientation of the surface. For example, the dissolution rate of (100) silicon in alkaline solutions can be many orders of magnitude larger than that of (111) substrate. A major cause of orientation dependent reactivity is the bonding condition of surface atoms. Characteristically, the silicon atoms on (111) surface have three bonds connecting to the substrate lattice while those on (100) surface have only two and are thus have weaker attachment to the surface compared to those on (111) surface.[99]

Also, a real surface has atomic structures associated with roughness and defects and thus the atoms at these lattice structures have different bonding conditions, some are similar to those on a (100) surface in terms of bonding characteristics; some are similar to a (111) surface and others are in between. Thus, a real surface may have varying degrees of reactivity determined by the concentration of the active atoms, which is a function of lattice structure determined by orientation, roughness, and type and density of defects.

The rate of an electrode reaction is a function of three principle types of species: charge carriers on the surface, active surface atoms and reactant species in the solution as illustrated in Figure 23. That is, $r \propto [h]*[Si_{active}]*[A]$. Carrier concentration and reactant concentration do not, in general, depend on surface orientation while active surface atoms may be a function of surface orientation. Anisotropic effect occurs when the rate determining step depends on the active surface atoms that vary with crystal orientation of the surface. On the other hand, reactions are isotropic when the concentration of active surface atoms is not a function of surface orientation or when the rate determining step does not involve active surface atoms.

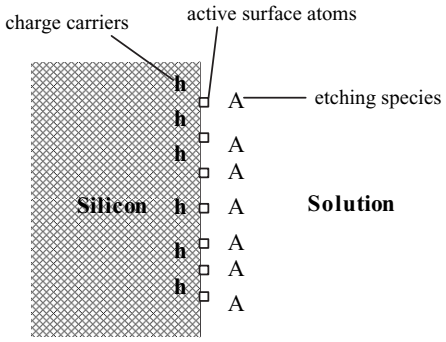


Figure 23. Schematic illustration of the elemental species involved in the dissolution reaction of silicon; h: carries inside the semiconductor, □: active silicon atoms an surface and A: etching species in the solution.

The surface, which is rough at the atomic scale, has no distinct crystallographic character and can be viewed as an amorphous surface as illustrated in Figure 24a. The density of active surface atoms on such surface is similar for different crystal orientations. Thus, when the number of atoms at kink sites, steps and other defects is close to the density of surface atoms, the surface loses its crystallographic nature. For surfaces of such a condition, the difference in reaction kinetics is small for different orientations. On the other hand, when the surface is less rough and has a clear crystallographic character, as shown in Figure 24b, the reaction on the surface is anisotropic when the rate determining step involves the surface atoms. When the surface is covered by an amorphous oxide film, the crystallographic nature of the substrate is masked and the reaction on such a surface is isotropic as shown in Figure 24c. This is the case for electropolishing of silicon electrodes in HF solutions at potentials higher than V_p where the dissolution is through the formation and dissolution of silicon oxide, which is amorphous in structure.

6. Reactions on the Surfaces of Silicon and Silicon Oxide

There are two basic reactions, that is, direct dissolution of silicon via reaction (I) and (II) and indirect dissolution through the formation and dissolution of silicon dioxide via reaction (II). The rates of both reactions increase with potential and proceed simultaneously on the surface of silicon. They compete with each other in rate and coverage of the surface area, and the relative surface coverage of the two reactions depends on the potential. At a low potential the surface coverage of oxide is zero when the rate of reaction (II), which is responsible for the formation of oxide, is low compared to the dissolution rate of the oxide. On the other hand, the coverage is one and the entire surface is covered with an oxide film when the rate of reaction (II) is much higher than the dissolution rate of oxide. Between the two extremes the surface is partially covered with oxide films at local places where the rate of reaction (II) equals the oxide dissolution rate. At low potentials the oxide formation rate is low compared to the dissolution rate and the surface is not covered by an oxide film. Thus, depending on the radius of curvature the bottom of a pore may be covered by an oxide film with varying thickness as shown in Figure 25.

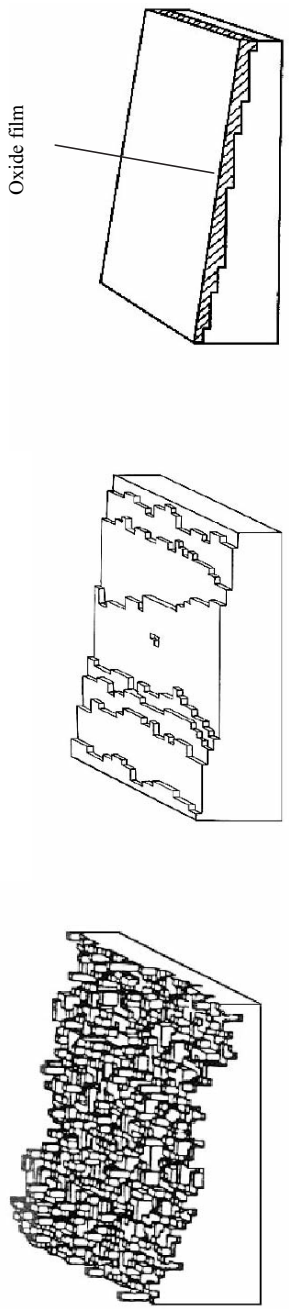


Figure 24. Schematic illustrations of the conditions of surface lattice structure: (a) amorphous-like surface with no identity of orientation, (b) surface with kinks, steps and terraces characteristic of certain crystalline orientation and (c) surface with no identity of the lattice structure of the crystal due to the coverage of an amorphous oxide film.

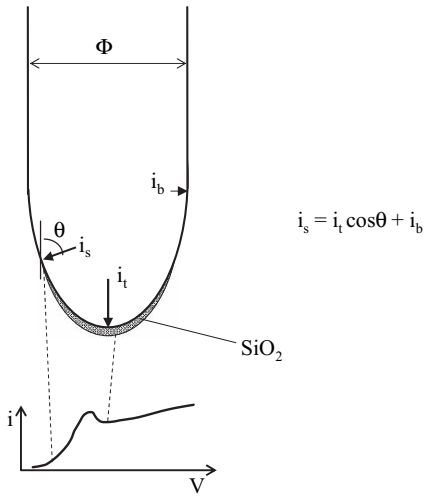


Figure 25. Schematic illustration of current variation and coverage of silicon oxide on the surface of a pore bottom.

Such reactions processes are responsible for the transition from PS formation to electropolishing with increasing potential as typically revealed in an i-V curve.[18] PS formation can occur when the surface is not or only partially covered by oxide. Once the whole surface is covered with an oxide film further reaction can only proceed through the formation of oxide followed by its dissolution. Further increasing the potential will only result in an increase of oxide film thickness. On the other hand, increasing HF concentration will increase the dissolution rate of oxide. The presence of oxide on the silicon surface in the PS formation region and its increase with potential have been experimentally observed.[98]

When the surface is completely covered by an oxide film, dissolution becomes independent of the geometric factors such as surface curvature and orientation, which are responsible for the formation and directional growth of pores. Fundamentally, unlike silicon, which does not have an atomic structure identical in different directions, anodic silicon oxides are amorphous in nature and thus have intrinsically identical structure in all orientations. Also, on the oxide covered surface the rate determining step is no longer electrochemical but the chemical dissolution of the oxide.[1]

7. Distribution of Reactions and their Rates on Pore Bottoms

For a stably growing PS the reactions and the rates are different on the pore walls and on the pore bottoms. Furthermore, they are different at different positions of a pore bottom due to the difference in the radius of curvature. The current is the largest at the pore tip because there the radius of curvature is the smallest. It decreases from the pore tip to the pore wall as radius of curvature increases. On the other hand, since the reactions depend on the current density, for a given condition, direct dissolution of silicon dominates at a relatively low current range while oxide formation and dissolution dominate at a higher current range. Thus, oxide formation and dissolution tend to occur at the pore tips at a lower potential than at the side of the pore bottom. There is a distribution of the kind of reactions along the pore bottom.

For a pore to propagate under a steady state the current density on the side of the pore bottom, i_s, and that at the pore tip, i_t, as illustrated in Figure 25 have the relation.

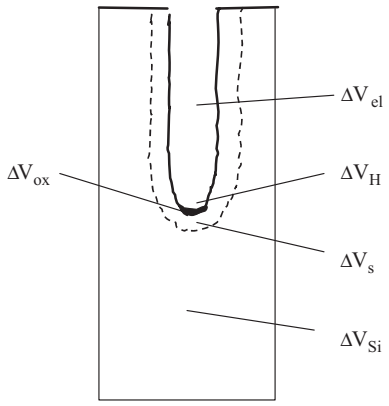
$$i_s = i_t \cos \theta + i_b \tag{5}$$

in which i_b is the extra current mostly due to the anisotropic effect and is responsible for the formation of side pores. The current density at different sites on the bottom depends on θ. It is the largest at the tip where $\theta = 0$, and is the smallest on the boundary of the bottom where $\theta = 90^o$. Such a distribution of current density is provided by the distribution of radius of curvature along the pore bottom, which determines the field at the silicon surface and the nature of the reactions. For different HF concentrations and

potentials the current density distribution on the pore bottom is different and so is the shape of the pore bottom. When the pore bottom is curved to such an extent that the current density distribution on the pore bottom for a given HF concentration and a given potential satisfies Eq. (5), pores will propagate stably.

(*i*) *Phases in Current Path*

As illustrated in Figure 26, which is a varied presentation for a single pore from the scheme shown in Figure 19, there are five possible phases in the current path in which significant potential drops may occur. The distribution of the applied potential in the different phases of the current path depends on doping type and concentration, HF concentration, current density, potential, illumination intensity and direction. The phases in the current path



$$\Delta V_{app} = \Delta V_{si} + \Delta V_s + \Delta V_{ox} + \Delta V_H + \Delta V_{el}$$

Figure 26. Potential drops along the current path in a pore; ΔV_{Si} = potential drop in silicon substrate, ΔV_s = potential drop in the space change layer ΔV_{ox} = potential drop in oxide ΔV_H = potential drop in the Helmholtz layer, ΔV_{el} = potential drop in electrolyte.

that take large portions of the change in the applied potential, have the consequence of changing current distribution along the pore bottoms, and thus affect the morphology of PS.

The resistance of the electrolyte causes a potential drop that is linearly distributed in the electrolyte inside the pores and thus does not have an effect on the current distribution on the pore bottom. However, the potential drop in the electrolyte has an important effect in maintaining the flat growth front of the PS layer as will be discussed later.

For moderately doped substrates, when the surface is free of oxide the change of potential is mostly dropped in the space charge layer and in the Helmholtz double layer. The reactions are very sensitive to geometric factors. The reaction that is kinetically limited by the processes in the space charge layer is sensitive to radius of curvature, while that limited by the processes in the Helmholtz layer is sensitive to the orientation of the surface. Depending on the relative effect of each layer the curvature effect versus anisotropic effect can vary.

When the pore bottom is covered by an oxide, the change of applied potential occurs almost completely in the oxide due to the very high resistance of the oxide. The rate of reactions is now limited by the chemical dissolution of the oxide on the oxide covered area. When the entire pore bottom is covered with an oxide the rate of reaction is the same on the entire surface of the pore bottom. As a result, the bottom flattens and the condition for PS formation disappears. The change of oxide coverage on the pore bottom can also occur when diffusion of the electrolyte inside deep pores becomes the rate limiting process. Since the current at which formation of an oxide occurs increases with HF concentration, a decreased HF concentration at pore bottom due to the diffusion effect can result in the formation of an oxide on the pore bottom of a deep pore at a condition that does not occur in shallow pores.

When the resistance of the substrate is high and a significant amount of potential is dropped in the substrate, the potential drop may not be uniform along a curved pore bottom due to the non-linear potential distribution on the material surrounding the bottom. Formation of macro PS on lowly doped materials becomes possible under such a condition.

8. Formation Mechanisms of Morphological Features

Formation of PS is due to preferential dissolution of a silicon surface; the rate is larger at some areas of the surface relative to others. Such relative rates with respect to the spatial position of the areas, on which these processes occur, are determined more by the relative nature and less by the absolute nature of the physical and chemical dimensions and events. Specifically, the relative nature of the following aspects is important in determining the morphology of PS:

1. Radius of curvature relative to the thickness of space charge layer
2. Relative depletion of carriers at pore tips to wall region
3. Relative dissolution rate on the (100) surface to that on the (111) surface
4. Density of surface atoms at steps and kinks relative to surface atomic density
5. Relative contribution of the preferential dissolution along direction of carrier source versus that along <100> direction
6. Relative growth rate of the side pores to that of the main pores
7. Change of profile of pore bottom due to random perturbations relative to pore size
8. Formation rate of oxide film relative to its dissolution rate
9. Rate of direct dissolution of silicon relative to indirect dissolution via oxide formation and dissolution
10. Relative share of potential change among the phases along the current path
11. Potential drop in electrolyte within a pore relative to those within the surrounding pores
12. Distance between surface and source of holes relative to thickness of space charge layer

(*i*) *Pore Diameter and Interpore Spacing*

Pore diameter and interpore spacing are determined by two large groups of factors: those that affect carrier density on the surface of a pore bottom and those that affect only the distribution

of the reactions. The first group of factors include doping type and concentration, potential, and illumination direction. The second group of factors include current density, HF concentration, and illumination frequency and intensity.

Except for the macro PS on lowly doped p-Si, both the pore diameter and wall thickness are largely determined by the thickness of the space charge layer. Thus, in general, pore diameter has the same order of magnitude as the thickness of space charge layer. Wall thickness is generally less than twice the space charge layer thickness. Because of the overlapping of the two space charge layers the wall region is depleted of carriers and is thus not conductive. If the wall thickness is larger than twice the space charge layer thickness the walls are not depleted of carriers and dissolution can still occur to form new pores on the walls. Under certain conditions such as backside illumination and surface patterning for n-Si, pore walls can be much thicker than twice the space charge layer thickness because of the favourable position of pore tips for the photo holes generated at the backside.

For a PS formed for controlled space charge layer thickness, the actual wall thickness depends on the relative dissolution rates between the edge of a pore bottom (see Figure 25) and the tip that is between i_b and i_t. If i_b is comparable to i_t, significant dissolution occurs at the edge of the pore bottom before the pore tip propagates far away. This will result in a thin wall or even no wall at all. On the other hand, if i_b is very small compared to i_t, the pore tip will propagate relatively fast so that, before much dissolution occurs on the edge of the pore bottom, the edge has already moved into the wall region where dissolution is virtually stopped due to lack of carriers. This will generate relatively thick walls.

At a given anodic current density, the potential is different for different doping types and concentrations due to different current conducting mechanisms. The combination of doping and potential thus provide a wide range of thickness of space charge layer and thus a wide range of pore diameters. For n-Si the space charge layer and thus the dimension of the pores increases with decreasing doping concentration and with increasing potential. For p-Si, on the other hand, increasing potential reduces the thickness of the space charge layer and results in smaller pores. However, when doping concentration is lower than a certain level,

the effect of substrate resistivity becomes important and formation of macro pores on p-Si can also occur. When this occurs, one obtains a PS with two distributions of pores with a micro PS layer on the surface of macro pores.

It needs to be pointed out that formation of macro pores can also occur on moderately or highly doped p-Si in the transition region (see Figure 4). Formation of macro pores on p-Si in the transition region was first noted on highly doped p-Si in aqueous HF solution[18] and was recently more systematically investigated in various non-aqueous solutions for moderately doped p-Si.[55] The macro pores formed in the transition region (near the limit of conditions for PS formation) may not be consistent with respect to size and depth from area to area as the coverage of PS may not be uniform on the surface. Also, the pores formed in this region generally have very thin walls relative to the size of pores. The formation of such macro pores can be viewed as a dimensional transition from pores of small radius of curvature in the exponential region to the flat surface that has an infinite radius of curvature. The formation of such large pores is related to the increased coverage of oxide on pore bottoms from the exponential region to electropolishing region. Since formation and dissolution of oxide strongly depend on the amount of water in the electrolyte, the condition for occurrence of the transition region and formation of macro pores on p-Si can greatly vary with type of solution as reported recently.[55]

Under front illumination on n-Si, photo carriers are generated from the surface to the depth of penetration determined by the wavelength of the light. The part of the space charge layer that has a field effect on the photo carriers varies from almost zero to the full thickness of the space charge layer. Correspondingly, the size of pores and wall thickness for the PS formed under illumination may vary from zero (i.e., corrosion of PS) to that comparable to the thickness of the space charge layer, resulting in a fractal-like PS structure.

When the current is larger than certain value, at which oxide starts to form at the tip of pores, increasing current density will increase the coverage of the oxide film on the pore bottom. In such a case the relative change in the current at the pore tip $\Delta i_t/i_t$ with an increase in the current in the pore, Δi, is less than that at the side of the pore bottom, $\Delta i_s/i_s$, that is, $\Delta i_t/i_t < \Delta i_s/i_s$. The

curvature of the pore bottom will increase resulting in larger pores and thinner walls. On the other hand, when the current is relatively small so that the current density at the pore tips is still much less than that required for the formation of oxide, an increase in applied current density will cause a relatively larger increase at the tip than on the side of the pore bottom (The slope of an i-V curve in this current range increases with increasing current density.), that is, $\Delta i_t/i_t > \Delta i_s/i_s$. As a result, the pore bottom becomes sharper with increasing current density. On the other hand, increasing the concentration of HF increases the dissolution rate of oxide, which in turn increases the critical current density at which the surface is covered by oxide. As a result, the pores become smaller and walls become thicker with increasing HF concentration.

(*ii*) *Initiation of Pores*

A silicon surface, no mater how well it is prepared, is not perfectly flat at the atomic scale, but has surface defects such as surface vacancies, steps, kinks sites, and dopant atoms. The dissolution of the surface is thus not uniform but modulated at the atomic scale with higher rates at the defects and depressed sites. The micro roughness of the surface will increase with the amount of dissolution due to the sensitivity of the reactions to surface curvature associated with the micro depressed sites. These sites, due to the higher dissolution rates, will evolve into pits and eventually into pores. Depending on the condition, a certain amount of dissolution is required before the initiation of pores on all types of materials.

The pores so initiated are very small in size but large in number due to the nature of surface defects. However, for the bulk PS formed under a steady state at a given anodization condition, the morphology has determined characteristics in terms of pore size, density, branching etc. Thus, the tiny pores initiated on the surface are not stable but tend to grow into the size required for such bulk PS. As the pores propagate into the bulk, some grow gradually in size while some vanish. The thickness of the initiation phase is comparable to the diameter of the pores grown at the steady state. Thus, the initiation layer is very thin for a micro PS and relatively thick for a macro PS.

Although surface defect sites are involved in the initiation of pores they do not determine the density and dimension of the pores in the bulk PS. The bulk morphology of PS is determined by the property of semiconductors and anodization conditions. However, under certain conditions, such as using surface patterning to generate initiation sites, the bulk PS morphology can be controlled to some extent.

(*iii*) *Flatness of the Growth Front of PS*

The growth front of a PS layer is generally flat and parallel to the initial surface of the substrate, meaning that the pores, which may have a range of diameters and shapes, propagate at essentially the same rate. This is due to the effect of electrolyte resistance within individual pores. The potential drop in the electrolyte within the pores, ΔV_{el}, plays the critical role of maintaining the same growth rate for different pores. For a pore of length, l, the potential drop in the electrolyte with a resistivity, ρ, inside the pore at a current density i is $\Delta V_{el} = i*\rho*l$. ΔV_{el} is independent of pore diameter and shape but depends on one geometric factor, that is, the length of the pore. Assume one pore is significantly ahead of the surrounding pores, i will increase due to the favourable position of the pore bottom. This will then increase the potential drop inside the pore, ΔV_{el}, which in turn will result in a reduction of the potential available for other phases in the current path. The reduction of the potential in other phases will then reduce the current density and slow down the growth of this particular pore. Thus, due to such a process individual pores can only grow at the same rate and, as a result, the growth front of a PS layer is flat.

(*iv*) *Orientation and Branching of Pores*

Orientation, branching and straightness of pores are related to similar factors, namely source of holes and anisotropic effect. Orientation of pores can be explained based on the current difference between (100) and (111) orientations, $\Delta i_{<100>} = i_{<100>} - i_{<111>,}$ and the difference between current densities on the pore bottoms of main pores and on side pores, $i_{hole,\ main}$ and $i_{hole,\ side}$, as shown in Figure 27. $i_{hole,\ side}$ is smaller than

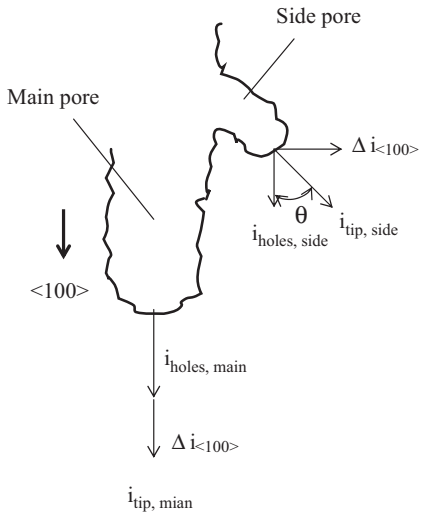


Figure 27. Current components at the tips of main pores and side pores.

$i_{hole,\ main}$ due to the unfavourable geometric location. At the tip of main pores, the total current density can be expressed as

$$i_{t,\ main} = i_{hole,\ main} + \Delta i_{<100>} \tag{6}$$

while at the tips of side pores the current density is

$$i_{t,\ side} = i_{hole,\ side}\cos\theta + \Delta i_{<100>}\sin\theta \tag{7}$$

Side pores grow at an angle, $\tan\theta = \Delta i_{<100>} / i_{hole,\ side}$, from the growth direction of the main pores. When the side pores become normal to the main pores when $\Delta i_{<100>} >> i_{hole,\ side}$, drendritic PS is formed. Also, when $i_{hole,\ main} >> i_{hole,\ side}$ the side pores do not grow very long before growth stops as the tip of the main pore quickly advances such that the material near the bottom of the side pores are quickly depleted of carriers. Little branching occurs under such a condition.

The stability of the profile of the pore bottom against local current perturbations determines the straightness of pores. Assume the amount of material dissolved, Δm, due to a current perturbation on a pore bottom, Δi, is $\Delta m = A\Delta i * t$ where t is time and A is a constant. This amount of dissolution may cause only a small change in the bottom profile of a large pore but can significantly change that of a small pore and may alter the growth direction of the small pore. This is why large pores are generally straighter than small pores because the bottom profile of large pores are more stable against random perturbations in the system.

(*v*) *Two-Layer PS*

Two-layer PS with a micro PS on top of a macro PS layer is formed on lowly doped p-Si or illuminated n-Si. For lowly doped p-Si two-layer PS can form when the conditions are such that the space charge layer and the resistive layer differ in dimension by several orders of magnitude.

Figure 28 illustrates schematically the two-layer PS formed on illuminated n-Si. The photo generated holes are located near the surface and flow to various directions depending on the direction of the field inside the walls of the micro PS (Note that the walls in these micro PS, in the order of nanometers, are much smaller than the thickness of space charge layer). These photo carriers can result in the dissolution of PS without the assistance of an applied anodic potential and are responsible for the etching of the PS layer. On the other hand, the photo carriers generated at different depths within the space charge layer are collected at the surface of the macro pore, resulting in the formation and growth of the micro PS on the surface of macro pore. The holes that are generated beyond the space charge layer are mostly collected and

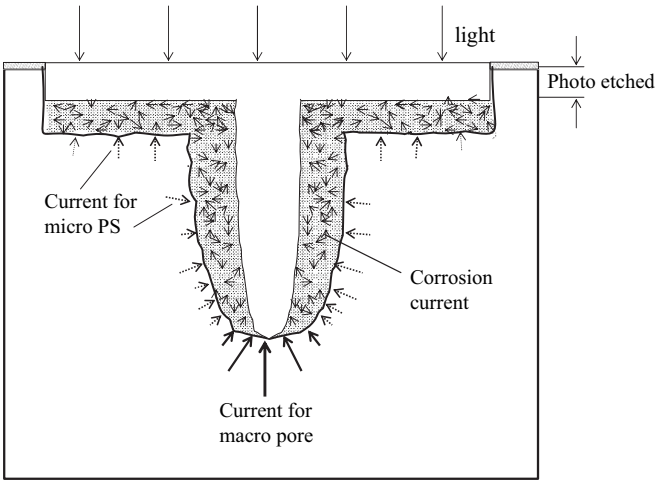


Figure 28. Schematic illustration of formation of a macro pore partially filled by a micro PS on n-Si under illumination.

react at the bottom of the macro pore resulting in the growth of the macro pore. Also, depending on the current density relative to that required for the formation of oxide, the tip area may or may not be covered with an oxide film. When the entire pore bottom is not covered with oxide, the macro pores will be fully filled with micro PS. On the other hand, when the tip area of pores is partially covered with an oxide film, macro pores are only partially filled with micro PS.

The formation of two-layer PS on p-Si involves two different physical layers in which the potential-current relations are sensitive to the radius of curvature. The space charge layer of p-Si under an anodic potential is thin, which is responsible for the formation of the micro PS. The non-linear resistive effect of the highly resistive substrate is responsible for that of macro PS. The effect of high substrate resistivity should also occur for lowly doped n-Si. However, under normal conditions, the thickness of the space charge layer under an anodic potential, at which macro PS is formed, is on the same order of magnitude as the dimension

of the resistive layer. The effect of these two different causes on the pore diameter is thus not normally distinguishable. If the conditions can be controlled such that the pores formed due to the effect of the space charge layer and those formed due to the effect of the resistive layer have a size difference by at least one order of magnitude it could be possible to obtain two-layer PS on n-Si.

For macro pores at a stable growth condition, the distribution of current density on the surface of an individual pore bottom is bell-like as shown in Figure 29a. For micro PS, the pores tend to grow in a zigzag fashion and the distribution of current at the dissolution front of PS is highly modulated across the surface as illustrated in Figure 29b. For two-layer PS, in which macro pores can be filled or partially filled with micro PS, the current profiles are illustrated in Figure 29c and 29d. The bell-shape distribution of the current responsible for growth of macro pores is consisted of fine modulations of the current responsible for growth of micro PS. When the current densities are such that no oxide occurs on any area of the pore bottom the macro pore is fully filled by micro PS. On the other hand, when the current density is such that oxide forms at the tip, the macro pore is only partially filled by micro PS.

(*vi*) *Dissolution of PS During its Formation*

The dissolution of PS during PS formation may occur in the dark or under illumination. Both are essentially corrosion processes, by which the silicon in the PS is oxidized and dissolved with simultaneous reduction of the oxidizing species in the solution. The material in the PS, which is distant from the growing front is little affected by the external bias due to the high resistivity of PS and is essentially at the open circuit potential (OCP). Such corrosion process is responsible for the formation of micro PS of certain thickness (stain film) in HF solutions containing oxidants under an unbiased condition.

Because the dissolution of silicon in HF solution requires holes, the corrosion rate of silicon in HF solutions at OCP is very low due to the unavailability of holes. But due to the large surface area of PS, the amount of dissolution still has a significant effect on the density of PS (e.g., PS density decreases with increasing PS thickness).

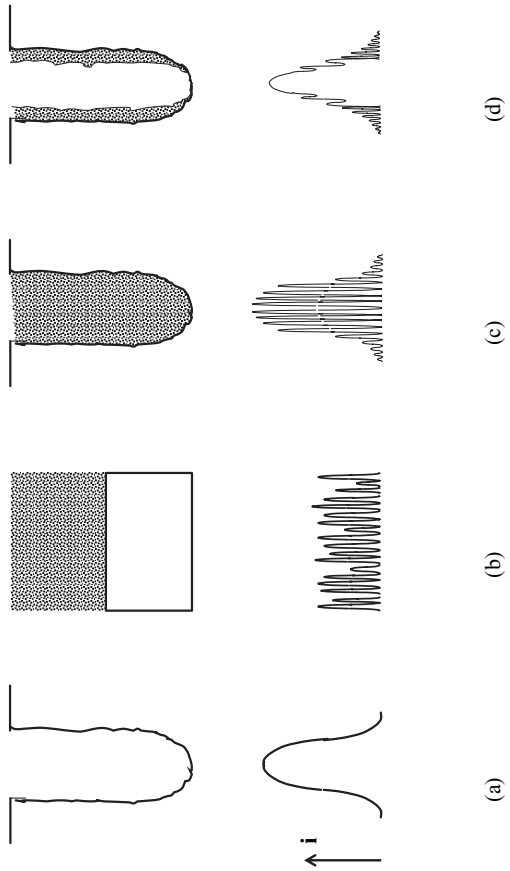


Figure 29. Schematic illustration for different types of PS and the current distributions; (a) void macro pore (b) micro PS (c) macro pore filled with micro PS (d) macro pore partially filled with micro PS.

Illumination generates holes within the material of PS and causes photo corrosion of PS that is much faster than that in the dark. Depending on illumination intensity and time, the pore walls in a PS can be thinned to various extents by the photo induced corrosion. This corrosion process is responsible for the etched crater between the initial surface and the surface of PS as illustrated in Figure 28. It is also responsible for the fractal structure of the micro PS formed under illumination.

V. SUMMARY

This chapter provided a conceptual analysis of the various aspects of morphology and formation mechanisms of porous silicon based on available information on the fundamental reaction processes on the silicon/electrolyte interface. According to the global mechanistic scheme described, the diverse types of PS can be categorized into three groups: 1) space charge layer controlled. This includes all PS except for the macro PS formed on p-Si. The diameter of the pores in these PS types is comparable to the width of space charge layer; 2) substrate resistance controlled. This includes the macro PS formed in lowly doped p-Si and possibly on lowly doped n-Si (a prediction); 3) photo carrier controlled. This includes all micro PS in single and two-layer PS formed under illumination.

Some morphological features of porous silicon in relation to formation condition as defined by the kinetics according to current and HF concentration are illustrated Figure 30. The lines defining the regions are determined by the nature of the reactions and are independent of the doping type and concentration (see Figure 4). One important factor governing the change of one region to the other is the formation and coverage of an anodic oxide film on the surface. The coverage of oxide on the surface increases from zero in PS region A to only at pore tips in PS region B to part of the surface in the transition region to full coverage in the electropolishing region.

Formation of porous silicon is an anodic dissolution process, which consists of carrier transport in the semiconductor, electrochemical reactions at the interface, and mass transport of the reactants and reaction products in the electrolyte. There are a

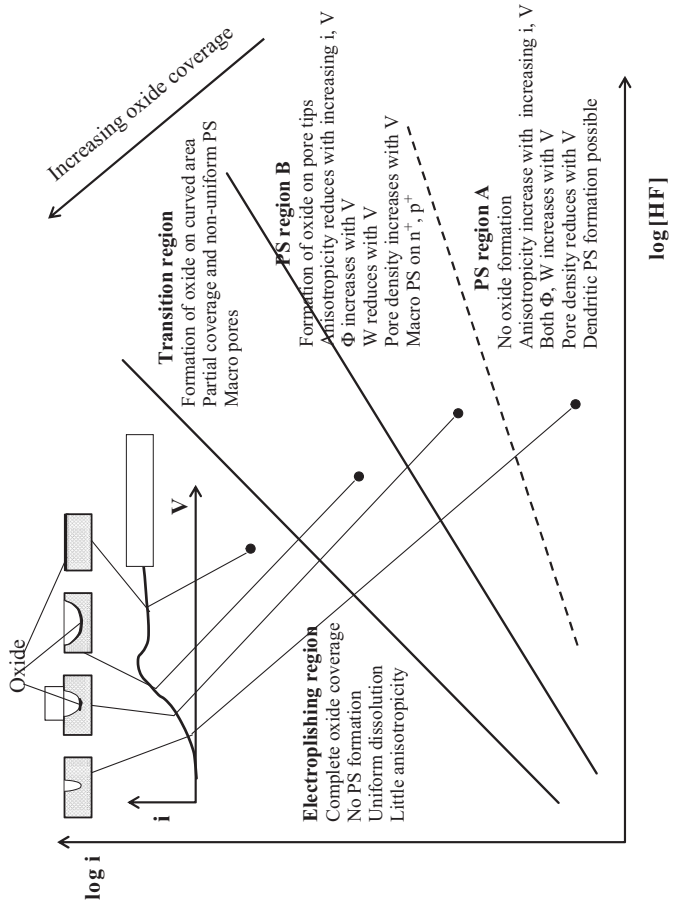


Figure 30. Conditions for the formation of various morphological features of PS.

number of reactions involved at the interface and these reactions consist of several steps and sub-reactions. At any given time, the dissolution kinetics can be controlled by any one or several of these steps. The distribution of reactions along a pore bottom under a steady state condition during pore propagation must be such that pore walls are relatively less active than the pore tip. Then, the dissolution reactions are concentrated at the pore tip to result in the preferential dissolution and the formation of pores. The formation of pores is the consequence of spatially and temporally distributed reactions.

The fundamental reason for the uneven distribution of reactions is that the rate of electrochemical reactions on a semiconductor is sensitive to the radius of curvature of the surface. This sensitivity can either be associated with the thickness of the space charge layer or the resistance of the substrate. Thus, when the rate of the dissolution reactions depends on the thickness of the space charge layer, formation of pores can in principle occur on a semiconductor electrode. The specific porous structures are determined by the spatial and temporal distributions of reactions and their rates which are affected by the geometric elements in the system. Because of the intricate relations among the kinetic factors and geometric elements, the detail features of PS morphology and the mechanisms for their formation are complex and greatly vary with experimental conditions.

REFERENCES

1. X. G. Zhang, *Electrochemistry of Silicon and Its Oxide*, Kluwer Academic, New York, 2001
2. D. R. Turner, *J. Electrochem. Soc.*,**105** (1958) 402.
3. A. Uhlir, Jr., *The Bell System Technical Journal*, **March** (1956) 333.
4. R. Memming and G. Schwandt, *Surf. Sci*. **4** (1966)109.
5. R. L. Meek, *J. Electrochem. Soc*.,**118** (1971) 437.
6. M. J. J. Theunissen, *J. Electrochem. Soc*., **119** (1972) 351.
7. Y. Arita and Y. Sunahama, *J. Electrochem. Soc.,* **124** (1977) 285.
8. Y. Watanabe, Y. Arita, T. Yokoyama and Y. Lgarashi, *J. Electrochem. Soc*., **122** (1975) 1315.
9. G. Bomchil, R. Herino, K. Barla and J. C. Pfister, *J. Electrochem. Soc*.,**130** (1983) 1611.
10 T. Uganami and M. Seki, *J. Electrochem. Soc*.,**125** (1978) 1339.
11. V. Labunov, I. Baranov and V. Bondarenko, *Thin Solid Film*, **64** (1979) 479.

12. M. I. J. Beale, J. D. Benjamin, M. J. Uren, N. G. Chew, and A. G. Cullis, *J. Crystal Growth*, **73** (1985) 622.
13. S. F. Chuang, S. D. Collins, and R. L. Smith, *Appl. Phy. Lett.,* **55**(7), 14 August, (1989) 675.
14. R. L. Smith and S. D. Collins, *J. Appl. Phy*. **71**(8), 15 April, (1992) R1.
15 R. L. Smith, S. F. Chuang, and S. D. Collins, *J. Electronic Materials*, **17** (1988) 228.
16. T. Unagami, *J. Electrochem. Soc*., **127** (1980) 476.
17 V. P. Parkhutik, L. K. Glinenko, and V.A. Labunov, *Surface Technology*, **20** (1983) 265.
18 X. G. Zhang, S. D. Collins, and R. L. Smith, *J. Electrochem. Soc.,***136** (1989) 1561.
19 V. Lehmann, *J. of Electroanalytical Chemistry*,**140** (1993) 2836.
20 V. Lehmann and H. Foll, *J. Electrochem. Soc*.**137** (1990) 653.
21 V. Lehmann, *Thin Solid Films*, **255 (**1995) 1.
22 V. Lehmann and U. Gosele, *Advanced Materials*, **4** (1992) 114.
23 V. Lehmann and U. Gosele, *Appl. Phys. Lett*., **58**(8), 25 February, (1991) 856.
24 X.G. Zhang, *J. Electrochem. Soc*., **138** (1991) 3750.
25 M. H. Al Rifai, M. Christophersen, S. Ottow, J. Carstensen, and H. Foll, *J. Electrochem. Soc*., **147** (2000) 627.
26 T. Osaka, K. Ogasawara, and S. Nakahara, *J. Elechochem. Soc*., **144** (1997) 3226.
27 C. Levy-Clement, A. Lagoubi and M. Tomkiewicz, *J. Electrochem. Soc*., **141** (1994) 958.
28 C. Levy-Clement, A. Lagoubi, R. Tenne and M. Neumann-Spallart, *Electrochimica Acta*, **37(**5) (1992) 877.
29 A. A. Yaron, S. Bastide, J. L. Maurice and C. L. Clément, *J. Luminescence*, **57** (1993) 67.
30 Y. Arita, *Journal of Crystal Growth*, **45** (1978) 383.
31 C. Levy-Clement, A. Lagoubi, and M. Tomkiewicz, *J. Electrochem. Soc*., **141** (1994) 958.
32 M. M. Rieger and P. A. Kahl, *J. Elechochem. Soc*., **142** (1995) 1490.
33 E. K. Propst and P. A. Kohl, *J. Electrochem. Soc.,* **141** (1994) 1006.
34 E. A. Ponomarev and C. Levy-Clement, *Electrochemical and Solid-State Letters*, **1** (1998) 42.
35 R. B. Wehrspohn, J. -N. Chazalviel, and F. Ozanam*, J.Electrochem. Soc.,* **145** (1998) 2958.
36 S. Cattarin, E. Pantano, and F. Decker, *Electrochemistry Communications*, **1** (1999) 483.
37 J. -N. Chazalviel, R. B. Wehrspohn, and F. Ozanam, *Materials Science and Engineering*, **B69-70** (2000) 1.
38 R. B. Wehrspohn, F. Ozanam, and J. -N. Chazalviel, *J. Electrochem. Soc*., **146** (9) (1999) 3309.
39 V. Lehmann and S. Ronnebeck, *J. Electrochm. Soc*., **146** (1999) 2968.
40 J. Carstensen, M. Christophersen, and H. Foll, *Materials Science and Engineering,* **B69-70** (2000) 23.
41 H. Foll, J. Carstensen, M.Christophersen, and G. Hasse, *Physica Status Solidi,* a(182) (2000) 45

42 J. Carstensen, M. Christophersen, and H. Foll, *Physica Status Solidi*, (2000) 63.
43 G. Hasse, M. Christophersen, J. Carstensen, and H. Foll, *Physica Status Solidi,* a(182) (2000) 23.
44 P. Jaguiro, S. La Monica, S. Lazarouk and A. Ferrari, *Electrochemical Society Proceedings*, **97-7** (1997) 296.
45 A. Valance, *Physical Review B,* **55** (1997) 9706.
46 S. La Monica, P. Jaguiro, and A. Ferrari*, Electrochemical Society Proceedings,* **97-7** (1997) 140.
47 R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*, John Wiley & Sons, New York, 1977.
48 I. Ronga, A. Bsiesy, F. Gaspard, F. Herino, M. Ligeon, and F. Muller*, J. Electrochem. Soc.*, **138 (**1991) 1403.
49 F. Gaspard, A. Bsiesy, M. Ligeon, F. Muller, and R. Herino*, J. Electrochem. Soc.*, **136** (1989) 3043.
50 P. C. Searson and X. G. Zhang*, J. Electrochem. Soc*., **137** (1990) 2539.
51 Y. Kang and J. Jorne, *J. Electrochem. Soc*., **144(**9) (1997) 3104.
52 M. J. Eddowes, *J. Electroanal. Chem.*, **280** (1990) 297.
53 V. A. Burrows, Y. J. Chabal, G. S. Higashi, K. Raghavachari and S. B. Christman, *Appl. Phys. Lett.,* **53**(11) (1988) 998.
54 C. M. Gronet, N. S. Lewis, G. Cogan, and J. Gibbons, *Proc. Natl. Acad. Sci. USA*, **80 (**1983) 1152.
55 S. Lust and C. Levy-Clement, *J. Electrochem. Soc*., **149** (2002) C338.
56 E. Peiner, and A. Schlachetzki, *J. Elechochem. Soc*., **139** (1992) 552.
57 J. Rappich and H.J. Lewerenz, *Thin Solid Film*, **276** (1996) 25.
58 J.-N. Chazalviel, M. Etman, and F. Ozanam, *J. Electroanal. Chem.,* **297** (1991) 533.
59 P.C. Searson, J.M. Macaulay, and Prokes, *J. Electrochem. Soc*., **139** (1992) 3373.
60 R. Herino, G. Romchil, K. Boala, and C. Bertrand, *J. Electrochem. Soc*., **134** (1987) 1994.
61 I. Ronga, A. Bsiesy, F. Gaspard, F. Herino, M. Ligeon, and F. Muller, *J. Electrochem. Soc*., **138** (1991) 1403.
62 V. V. Doan and M. Sailor, *Science*, **256** (1992) 1791.
63 M. Christophersen, J. Carstensen, A. Feuerhake, and H. Foll, *Materials Science and Engineering*, **B69-70** (2000) 194.
64 L. Koker and K. W. Kolasinski, *Materials Science and Engineering*, **B69-70** (2000) 132.
65 X. G. Zhang, unpublished results.
66 R. C. Frye, *Mat. Res. Soc. Symp. Proc*., **33** (1984) 53.
67 M. Guendouz, P. Joubert, and M. Sarret, *Materials Science and Engineering*, **B69-70** (2000) 43.
68 H. Unno, K. Imai, and S. Muramoto, *J. Electrochem. Soc*., **134** (1987) 645.
69 H. Braumgart, R. C. Frye, F. Philipp, and H. J. Leamy, *Mat. Res. Soc. Symp. Proc*. **33** (1984) 63.
70 O. Teschke, *Appl. Phys. Lett*., **64(**15) (1994) 1986.
71 M. I. J. Beale, N. G. Chew, M. J. Uren, A. G. Cullis, and J. D. Benjamin, *Appl. Phys. Lett.,* **46**(1) (1985) 1095.
72 H. Sugiyama and O. Nittono, *Journal of Crystal Growth*, **103** (1990) 156.

73 K. D. Legg, A. B. Ellis, J. M. Bolts, and M. S. Wrighton, *Proc. Natl. Acad. Sci. USA*, **74(**10) (1977) 4116.

74 N. Noguchi, I. Suemune, M. Yamanishi, G. C. Hua, and Nobuo Otsuka, *Jpn. J. Appl. Phys.*, **31** (1992) L490.

75 M. Christophersen, J. Carstensen, and H. Foll, *Physica Status Solidi*, (a) 182, 1 (2000) 45.

76 S. -F. Chuang, S. D. Collins, and R. L. Smith, *Appl. Phys. Lett.,* **55**(15) (1989) 1540.

77 P. C. Searson, J. M. Macaulay, and F. M. Ross, *J. Aapl. Phys*., **72**(1) (1992) 253.

78 T. Osaka, K. Ogasawara, M. Katsunuma, and T. Momma, *J. of Electroanalytical Chemistry*, **396 (**1995) 69.

79 V. Labunov, V. Bondarenko, L. Glinenko, A. Dorofeev and L. Tabulina, *Thin Solid Films*, **137** (1986) 123.

80 V. Lehmann, R. Stengl, and A. Luigart, *Materials Science and Engineering*, **B69** (2000) 11.

81 V. Lehmann, A. Luigart and V. Corbel, *Electrochemical Society Proceedings*, **97-7** (1997) 132.

82 C. Gui, M. Elwenspoek, J. G. E., Gardeniers, and P. V. Lambeck, *J. Electrochem. Soc.,* **145** (1998) 2204.

83 E.A. Ponomarev and C. Levy-Clement, *Electrochemical Society Proceedings*, **97-7** (1997) 319.

84 C. Jäger, B. Finkenberger, W. Jäger, M. Christophersen, J. Carstensen, and H. Föll, *Materials Science and Engineering*, **B69-70** (2000) 199.

85 V. Lehmann and U. Grüning, *Thin Solid Films*, **297** (1997) 13.

86 J. E. A. M. van den Meerakker, R. J. G. Elfrink, F. Roozeboom and J. F. C. M. Verhoeven, *J. Electrochem. Soc*., **147**(7) (2000) 2757.

87 P. Allongue, C. H. de Villeneuve, M. C. Bernard, J. E. Peou, A. Boutry-Forveille, and C.Levy-Clement, *Thin Solid Films*, **297** (1997) 1.

88 St. Frohnhoff, M. Marso, M. G. Berger, M. Thonossen, H. Luth, and H. Munder, *J. Elechochem. Soc*., **142** (1995) 615.

89 M. Binder, T. Edelmann, T. H. Metzger, G. Mauckner, G. Goerigk, and J. Peisl, *Thin Solid Film*, **276** (1996) 65.

90 S. Billat, M. Thonissen, R. Arens-Fischer, M. G. Berger, M. Kruger, and H. Luth, *Thin Solid Films*, **297** (1997) 22.

91 R. Houbertz, U. Memmert, and R. J. Behm, *J. Vac. Sci. Technol*. *B*, **12(**6) (1994) 3145.

92 M. Thönissen, M. G. Berger, R. Arens-Fischer, O. Glüger, and H. Lüth, *Thin Solid Films*, **276** (1996) 21.

93 M. Thönisson, M. G. Berger, R. Arens-Fisher, O. Glück, M. Krüger, and H. Lüth, *Thin Solid Film*, **276** (1996) 21.

94 S. F. Chuang, S. D. Collins and R. L. Smith, *The Technical Digest of the Solid State Sensor and Actuator Workshop,* Hilton Head Island, SC, June 6-9, IEEE, New York, 1988, p.151.

95 S. Ronnebeck, J. Carstensen, S. Ottow, and H. Foll, *Electrochemical and Solid State Letters*, **2** (2000) 126.

96 S. Bengtsson and L. Engstrom*, J. Appl. Phys*., **66**(3) (1989) 1231.

97 A. G. Cullis and L. T. Canham, *Nature*, **353** (1991) 353.

98 A. Belaidi, M. Safi, F. Ozanam, J. N. Chazalviel, and O. Gorochov, *J. Electrochem. Soc.*, **46 (**1999) 2659.

3

Modeling Electrochemical Phenomena via Markov Chains and Processes

Thomas Z. Fahidy

Department of Chemical Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

I. INTRODUCTION

Named after the Russian mathematician Andrei Andreievich Markov (1856–1922), Markov chains and processes express the probability evolution of events, whose future, loosely speaking, is independent of their past. The terminological difference between the two designations is that Markov chains are related to discrete events, whereas Markov processes are related to continuous events. Put simply, Markov chains are discrete Markov processes. Both have gained prominence in physics, engineering, management sciences, game theory, etc., and, in general, in the study of complex processes. A comprehensive survey of historical background and development, literature and scope of applications by Tamir[1] is highly recommended to the seriously interested reader (the contents of the book transcend its title!). Fundamental theory has been presented in the classic texts e.g., by Pugachev,[2] Parzen,[3] Stark and Woods,[4] Rozanov,[5] and Sveshnikov.[6] Matrix techniques pertinent to Markov chains have been discussed e.g.,

Modern Aspects of Electrochemistry, Number 39, edited by C. Vayenas et al., Springer, New York, 2005.

by Gantmacher,[7] and Greenberg.[8] References [9-14] form a small, but illustrative sample of the continuing wide interest in the subject matter (also manifested by the 392 references listed on the Internet in early May 2004 under the title of Markov processes).

In the domain of electrochemistry and electrochemical engineering, the penetration of Markov theory has so far been rather scant. In an early exploratory study the usefulness of the Erlang formula (Section II.3) was demonstrated for the analysis of magnetically enhanced natural convection in electrolysis.[15] The interpretation of spatio–temporal discharge patterns as a probabilistic function of a Markov process indicates that the latter is a powerful tool in the analysis of multi–electrode data related to neuronal spike activity.[16] In an investigation of thin film growth processes dependent on the substrate interaction potential, Oprişan et al.[17] used a local Markov chain–based random function model, where the transition probability matrix describes Brownian motion in a rectangular lattice in order to produce 2D fractal patterns. Markov process–based random walks within a lattice was also used in the simulation of Brownian motion of ions in a force field,[18] for studying the effect of ionic migration and diffusion on the roughness of electrodeposits: the eventual disappearance of the influence of initial roughness is a concomitant conclusion. Random walk/Markov process models were also applied to diffusion-limited aggregates and electro-deposition clusters,[19-21] diffusion across ion-exchange membranes,[22] and to tank electrolyzers.[23]

In addition to providing a survey of previous studies, it is also the purpose of this chapter is to explore new avenues of potential applications to various electrochemical phenomena. A condensed theoretical background is first presented, and followed by a Markovian treatment of selected processes and phenomena of interest to electrochemists, electrochemical engineers, electrochemical technologists, and technical managers involved with electrochemical plants. Multidimensional analysis is excluded, however, from this essentially *whetting the appetite* exercise. Material in Section V relies heavily on author's personal (and hitherto unpublished) research. The presentation, although by no means exhaustive, intends to focus on basic concepts, potential utility and new horizons, rather than to pursue completeness.

II. A CONCISE THEORY OF MARKOV CHAINS AND PROCESSES

1. Fundamental Notions[5,6] of Markov chains

A physical system, occupying states s_0, s_1, …changes randomly its state from an initial state at zero time, i.e., a random variable $X(t)$ of the state undergoes consecutive transitions $X(0) \rightarrow X(1) \rightarrow X(2) \rightarrow \ldots$ The state is at s_i ($t = 0$) with initial probability $p_i^0 = \Pr[X(0) = s_i]$; $i = 1,2,\ldots$ Then, the transition probability

$$p_{ij} = \Pr[X(n+1) = s_j / X(n) = s_i] \; ; \; i = 1,2,... \tag{1}$$

yields the probability of the system being at state s_j at time or stage (n + 1), if it was at state s_i at time or stage (n). This probability does *not* depend on time (n). The random process $X(t)$ is called a Markov chain with stationary random probabilities, or simply a Markov chain. It follows that the probability $p_j(n)$ that the system will be in the state s_j at time (n) (or, equivalently, after n steps) may be expressed as

$$p_j(0) = p_j^0 \tag{2a}$$

$$p_j(n) = \sum_k p_k(n-1) p_{kj} \; ; \; n = 1,2... \tag{2b}$$

If, at $t = 0$ the system is in the state s_i, then $p_i^0 = 1$, $p_k^0 = 0$ for $k \neq i$, hence

$$\begin{aligned} p_{ij}(0) &= 1 \; ; \; j = i \\ p_{ij}(0) &= 0 \; ; \; j \neq i \end{aligned} \tag{3a}$$

and

$$p_{ij}(n) = \sum_k p_{ik}(n-1) p_{kj} \; ; \; n = 1,2... \tag{3b}$$

is the probability that the system will go from state s_i to state s_j in n steps. The elements p_{11}, p_{12}, p_{13}…form the first row, p_{21}, p_{22}, p_{23}…form the second row, etc. of the Markov matrix $\boldsymbol{P}$ with the important characteristic of

$$\boldsymbol{P}(n) = (\boldsymbol{P})^{n}\,;\ \boldsymbol{P}(0) = \boldsymbol{I} \tag{4}$$

where $\boldsymbol{I}$ is the identity matrix. The one-dimensional random walk is a classical Markov chain.

Working with Markov chains, confusion is bound to arise if the indices of the Markov matrix are handled without care. As stated lucidly in an excellent elementary textbook devoted to finite mathematics,[24] transition probability matrices must obey the constraints of a stochastic matrix. Namely that they have to be square, each element has to be non-negative, and the sum of each column must be unity. In this respect, and in order to conform with standard rules vector-matrix multiplication, it is preferable to interpret the probability p_{ij} as the probability of transition from state s_j to state s_i (this interpretation stipulates the standard $\boldsymbol{Pp}$ format instead of the $\boldsymbol{p}^{T}\boldsymbol{P}$ format, the latter convenient for the alternative $s_i \rightarrow s_j$ interpretation in defining p_{ij}).[5,6]

If $p_j(n) \rightarrow p_j^{*}$ as $n \rightarrow \infty$, then p_j^{*}; $j = 1,2,\ldots$ describe the final probability distribution of the system being at state s_j. Since as $n \rightarrow \infty$, $\boldsymbol{p}_j(n) = \boldsymbol{P}\boldsymbol{p}_j(n-1)$, the limiting probabilities are defined also by the classical eigenvalue theorem[7] $\boldsymbol{p}_j^{*} = \boldsymbol{P}\boldsymbol{p}_j^{*}$ with an implied eigenvalue of unity (Appendix), and the sum of all probabilities being equal to unity. A tutorial treatment of chemical kinetics[25] provides a simple illustration.

Similar to the discrete case, $p_i^{0} = \Pr[X(0) = s_i]$, $i = 1,2,\ldots$is the initial probability of the random process being in state s_i. The transition probability

$$p_{ij}(t) = \Pr[X(\tau + t) = s_j \,/\, X(\tau) = s_i]\ ;\ i, j = 1,2... \tag{5}$$

expresses the probability that the system is in state s_j after a time t, if it was in state s_i at time τ; this probability is independent of τ. If $p_{ij}(0) = 1$ for $j = i$, and zero for any $j \neq i$, then

$$p_{ij}(\tau + t) = \sum_k p_{ik}(\tau) p_{kj}(t) \ ; \ i, j = 1,2,... \tag{6}$$

for any arbitrary τ and t, and

$$p_j(t) = \sum_k p_k(\tau) p_{kj}(t); j = 1,2,... \ ; \ p_j(0) = p_j^0 \tag{7}$$

An important case is where the transitional probabilities can be expressed as $p_{ii}(\Delta t) = 1 - \lambda_i \Delta t$ and $p_{ij}(\Delta t) = \lambda_{ij}\Delta(t); j \neq i; i = 1,2,\ldots$ with λ_i being the density parameter of transition out of the state s_i, and λ_{ij} the density parameter of transition out of the state s_i to the state s_j . Defining $\lambda_{ii} = -\lambda_i$, $i = 1,2,\ldots$, the transition probabilities satisfy the forward Kolmogorov (also known as Chapman-Kolmogorov) equation

$$dp_{ij}(t)/dt = \sum_k \lambda_{kj} p_{ik}(t) \ ; \ i, j = 1,2... \tag{8}$$

and the backward Kolmogorov equation

$$dp_{ij}(t)/dt = \sum_k \lambda_{ik} p_{kj}(t) \ ; \ i, j = 1,2,... \tag{9}$$

with initial conditions $p_{ij}(0) = 1$, if $j = i$; zero, if $j \neq i$. The limiting probabilities are similar to the Markov chain case in the sense that

$$\lim p_j(t) = p_j^* \ ; \ t \to \infty \tag{10}$$

is the probability of the system occupying the state s_j at very large times, independent of the initial probability distribution.

Finally, if a Markov process is in some state s at some (reference) time t_0, and if θ is the random time required for the process to leave state s to go to some other state, then the probability that θ is larger than some arbitrary time t is given by

$$\Pr[\theta > t] = \exp(-\lambda t) \ ; \ t \geq 0 \tag{11}$$

where λ is a non-negative constant. This important theorem describes the conduct of a service system with exponential holding times in queuing theory.

2. The Erlang Formula

Taking the limit in Eq. (7) as $\tau \to \infty$, and considering the (more practical) case of a finite number of states m,

$$p_j^* = \sum_i p_i^* p_{ij}(t) \ ; \ j = 1,...,m \tag{12}$$

$$\sum_i p_i^* \lambda_{ij} = 0 \ ; \ j = 1,...,m \tag{13}$$

If there exist only two states s_0 and s_1 , it follows from Eq. (13) that $p_0^* \lambda_{01} + p_1^* \lambda_{11} = 0$. It is useful to invoke certain concepts of queuing theory at this stage, and consider the states as number of *servers* handling incoming *calls*, the arrivals of calls having a Poisson distribution with parameter λ. Then, $\lambda_{01} = \lambda$, and $\lambda_{11} = -\mu$, where μ is the density parameter of service. Eq. (13) can further be written as $\lambda p_0^* = \mu p_1^*$. If there are two servers (or three states s_0, s_1, and s_2 where s_0 represents the state of no server being engaged), Eq. (13) can be written as $p_0^* \lambda_{01} + p_1^* \lambda_{11} + p_2^* \lambda_{21} = 0$. Given that $\lambda_{01} = \lambda$; $\lambda_{11} = -(\lambda + \mu)$, and $\lambda_{21} = 2\mu$, the final form is $(\lambda + \mu) p_1^* = \lambda p_0^* + 2\mu p_2^*$. In general, the m-server pattern can be written as

$$\lambda p_0^* = \mu p_1^* \tag{14a}$$

$$(\lambda + j\mu) p_j^* = \lambda p_{j-1}^* + (j+1)\mu p_{j+1}^* \ ; \ j = 1,2,m-1 \tag{14b}$$

$$\lambda p_{m-1}^* = m\mu p_m^* \tag{14c}$$

From this set, the individual limiting probabilities are obtained by successive elimination as $p_1^* = (\lambda/\mu) p_0^*; \ p_2^* = (\lambda/\mu)^2 p_0^*/2;$

$...p_j^* = (\lambda/\mu)^j p_0^*/j!$ and taking into account that all probabilities must add up to unity, the Erlang formula

$$p_j^* = [(\lambda/\mu)^j/j!]/\sum_{j=0}^{m}(\lambda/\mu)^j/j! \tag{15}$$

is finally obtained. Several applications of this important expression to electrochemical systems will be illustrated.

3. Service System with Exponential Holding Time

If the incoming *call* distribution is Poisson, and the time to service a call has an exponential distribution (Eq. 11 with density parameter μ), then the transitional probabilities related to states s_0 (the server is free) and s_1 (the server is engaged) are $p_{10} = 1 - \exp(-\mu t)$; $p_{01} = 1 - p_{00}$; $p_{11} = 1 - p_{10}$; p_{00} and p_{11} are found by solving the Kolmogorov equations (Eq. 8)

$$dp_{00}/dt + (\lambda+\mu)p_{00} = \mu \tag{16}$$

$$dp_{11}/dt + (\lambda+\mu)p_{11} = \lambda \tag{17}$$

with initial conditions $p_{00} = p_{11} = 1$. Employing e.g., classical techniques, these linear ordinary differential equations are solved to

$$p_{00} = [r/(r+1)]\exp[-(\lambda+\mu)t] + 1/(r+1) \tag{18}$$

$$p_{11} = [1-r/(r+1)]\exp[-(\lambda+\mu)t] + r/(r+1) \tag{19}$$

where the $r = \lambda/\mu$ ratio portrays the service efficiency. Further,

$$p_{01} = 1-[1-r/(r+1)]\exp[-(\lambda+\mu)t] - 1/(r+1) \tag{20}$$

$$p_{10} = 1-[1-1/(r+1)]\exp[-(\lambda+\mu)t] - r/(r+1) \tag{21}$$

Consequently, the limiting probabilities are $p_{00}^{*} = 1/(r + 1)$; $p_{11}^{*} = r/(r + 1)$; $p_{01}^{*} = r/(r + 1)$ and $p_{10}^{*} = 1/(1 + r)$. If service is very efficient, i.e., $r \to 0$, s_0 is the dominating state (i.e., the server is mostly free), and in the opposite case when $r \to \infty$, s_1 is the dominating state (i.e., the server is mostly engaged).

III. MARKOVIAN MODELING OF ELECTRODE SURFACE PROCESSES

1. Random Walk of an Ion Between Adjacent Active Centres

In this relatively simple random walk model an ion (e.g., a cation) can move freely between two adjacent active centres on an electrode (e.g., cathode) with an equal probability ½. The centres are separated by L characteristic length units. When the ion arrives at one of the centres, it will react (e.g., undergoes a cathodic reaction) and the random walk is terminated. The centres are, therefore absorbing states. For the sake of illustration, $L = 4$ is postulated, i.e., s_1 and s_5 are the absorbing states, if 1 and 5 denote the positions of the active centres on the surface, and s_2, s_3, and s_4 are intermediate states, or ion positions, $L/4$ characteristic units apart. The transitional probabilities $p_{ij}(n) = \Pr[s_i \to s_j \text{ in } n \text{ steps}]$ must add up to unity, but their individual values can be any number on the [0, 1] domain.

In the simplest case, the random walk is equi-probable between two adjacent positions, but no *jumpover* is allowed, i.e., $p_{31} = p_{41} = \ldots = 0$ and the ion cannot be stagnant, i.e., $p_{22} = p_{33} = p_{44} = 0$ (however, $p_{11} = p_{55} = 1$). Table 1 indicates that the ion eventually reacts at one of the active centres, the probability of reaction at a centre depending on the initial position of the ion. If, e.g., the initial position was s_4 , the probability of the ion arriving eventually at the active centre s_1 is ¼, and at active centre s_5, ¾. The final state is reached essentially after about 25 steps (stages).

Table 2 shows the evolution of the ionic random walk with unequal transition probabilities, movement to active centre s_1 being twice more probable than to active centre s_5 .

Table 1
Elements of the Transition Probability Matrix of a Unidimensional Equiprobable Random Walk Model for an Ion Moving on an Electrode Surface; S_1 And S_5 are Absorbing States

n = 1	i = 1	i = 2	i = 3	i = 4	i = 5
j = 1	1	0	0	0	0
j = 2	½	0	½	0	0
j = 3	0	½	0	½	0
j = 4	0	0	½	0	½
j = 5	0	0	0	0	1

n = 10	i = 1	i = 2	i = 3	i = 4	i = 5
j = 1	1	0	0	0	0
j = 2	0.7344	0.0156	0	0.0156	0.2344
j = 3	0.4844	0	0.0312	0	0.4844
j = 4	0.2344	0.0156	0	0.0156	0.7344
j = 5	0	0	0	0	1

n ≥ 25	i = 1	i = 2	i = 3	i = 4	i = 5
j = 1	1	0	0	0	0
j = 2	¾	0	0	0	¼
j = 3	½	0	0	0	½
j = 4	¼	0	0	0	¾
j = 5	0	0	0	0	1

2. Markovian Analysis of Layer Thickness on a Surface

If the service system concept is adapted to a reactive solid surface, its active centres can be divided into a number of agglomerates or clusters, and the probability of N clusters being engaged can be related to the efficiency of the surface via the Erlang formula. The approach is illustrated by an oxide-thickness-on-wafers problem[26] where seven random measurements of oxide layer thickness obtained under identical experimental conditions are given as 1264, 1280, 1301, 1300, 1292, 1307, 1275 in Å units. The model postulates a single- cluster activity responsible for oxide layer thickness between 1260–1270; two-cluster activity for thickness between 1270–1280, etc…assigning a total of five clusters to the entire 1260–1310 range. Alternatively, activity levels labeled 0, 1, 2, … can be envisaged for each layer thickness range. Following elementary rules of statistical data classification (described in any

Table 2
Elements of the Transition Probability Matrix of a Unidimensional Unequiprobable Random Walk Model for an Ion Moving on an Electrode Surface; S_1 and S_5 are Absorbing States

n = 1	i = 1	i = 2	i = 3	i = 4	i = 5
j = 1	1	0	0	0	0
j = 2	2/3	0	1/3	0	0
j = 3	0	2/3	0	1/3	0
j = 4	0	0	2/3	0	1/3
j = 5	0	0	0	0	1

n = 10	i = 1	i = 2	i = 3	i = 4	i = 5
j = 1	1	0	0	0	0
j = 2	0.9229	0.00867	0	0.00433	0.0641
j = 3	0.7861	0	0.0174	0	0.1965
j = 4	0.5125	0.0173	0	0.00867	0.4615
j = 5	0	0	0	0	1

n ≥ 25	i = 1	i = 2	i = 3	i = 4	i = 5
j = 1	1	0	0	0	0
j = 2	0.9333	0	0	0	0.0667
j = 3	4/5	0	0	0	1/5
j = 4	0.5333	0	0	0	0.4667
j = 5	0	0	0	0	1

text on elementary statistics) the cumulative distribution function is approximated by assigning cumulative frequencies to their respective class mid-points, and joining these values by a continuous curve. Cluster efficiency, defined as r = λ/μ, is computed via the Erlang formula (Eq. 15), expressed as

$$\Pr[N] = \left(r^N / N!\right) \Big/ \left(\sum_{j=0}^{N} r^j / j! \right) \tag{22}$$

where Pr[N] is obtained as the difference between the values of the cumulative distribution function evaluated at adjacent class mid-points. Table 3 demonstrates that the thickness level can be interpreted in terms of service efficiency, i. e. the faster the surface activity, the thicker the deposit layer on it. In queuing

Table 3
Cluster Efficiency by Erlang's Formula Applied to the Oxide Layer Thickness Problem[25]

Cluster no. N	1	2	3	4	5
Class interval	1260–1270	1270–1280	1280–1290	1290–3000	3000–3100
Pr[N]	0.22	0.15	0.14	0.27	0.22
r = λ/μ	0.28	0.80	1.54	3.59	4.24

theory, an equivalent statement would be that the faster the service (or the larger the number of servers available), the larger the number of clients that can be served.

3. Markovian Interpretation of an Anodic Deposit Formation ↔ Deposit Dissolution Process

Consider an electrode process consisting of the following (periodic) sequence: an anion is subjected to an anodic reaction producing a solid deposit on the anode, followed by partial desorption/re-dissolution from the surface into the electrolyte. In a Markovian framework, if s_1 is the ionic state, and s_2 is the surface deposit state, the transition probabilities p_{11}, p_{12}, p_{21}, p_{22} define the probability of the ion remaining in the ionic state, the deposit desorbing and re-dissolving into the ionic state, the ion undergoing reaction and forming a deposit, the deposit desorbing and re-dissolving, the deposit remaining on the surface, respectively, the indexing according to statements following Eq. (4).

The system is mathematically trivial if the transitional probabilities are equally ½, inasmuch as $\boldsymbol{P} = \boldsymbol{P}^{\mathbf{n}}$, regardless of the magnitude of n. This would be the case of sustained oscillation between ionic and deposit state. Table 4 demonstrates the evolution of the Markov chain in the case of unequal transition probabilities, in two specific instances. The speed of convergence to limiting values decreases with the degree of asymmetry in the P-matrix elements. The limiting values do not depend on the initial state, inasmuch as they are eigenvector elements related to the eigenvalue of unity. The complementary eigenvalues of 0 in the first, and 0.4 in the second instance have no importance, since

Table 4
Markov Chain Evolution for the Anodic Deposit Formation↔Deposit Dissolution Process. The Initial State is Purely Ionic ($p_1^0 = 1$; $p_2^0 = 0$)

n	p_1	p_2	p_1	p_2
	$p_{11} = 0.6$ $p_{12} = 0.7$ $p_{21} = 0.4$ $p_{22} = 0.3$		$p_{11} = 0.6$ $p_{12} = 0.2$ $p_{21} = 0.4$ $p_{22} = 0.8$	
1	0.6000	0.4000	0.6000	0.4000
2	0.6400	0.3600	0.4400	0.5600
3	0.6360	0.3640	0.3760	0.6240
4	0.6364	0.3636	0.3504	0.6496
5	0.6364	0.3636	0.3401	0.6598
∞	0.6364	0.3636	0.3333	0.6667

the former implies zero probabilities, and the latter implies a negative probability.

A three-dimensional variation of the theme is offered by the oscillatory behaviour of anodic copper dissolution into a NaCl/KSCN electrolyte mixture.[27] This is a complex process involving solid states CuSCN (pK_{sp} = 14.32) and Cu_2O or CuOH (pK_{sp} = 14), and CuCl (pK_{sp} = 5.92), and ionic species Cu^+, $CuCl_2^{2-}$, $CuCl_3^-$, and Cl^-. Among other plausible schemes,

Table 5
Markov Chain Evolution for the Oscillatory Anodic Dissolution of Copper.[26] $p_{11} = 0.1$; $p_{12} = 0.5$; $p_{13} = 0.4$; $p_{21} = 0.8$; $p_{22} = 0.1$; $p_{23} = 0$; $p_{31} = 0.1$; $p_{32} = 0.4$; $p_{33} = 0.6$; 1:Cl^-; 2:$CuCl_2^-$; 3:$CuCl_3^{2-}$; $p_1^0 = 1$; $p_2^0 = p_3^0 = 0$

n	p_1	p_2	p_3
1	1/10	8/10	1/10
2	0.4500	0.1600	0.3900
3	0.2810	0.3760	0.3430
4	0.3533	0.2624	0.3843
5	0.3203	0.3089	0.3708
10	0.3304	0.2933	0.3762
50	0.3303	0.2936	0.3761
∞	0.3303	0.2936	0.3761

transition from Cl^- to $CuCl_2^-$ and $CuCl_3^{2-}$, from $CuCl_2^-$ to Cl^- and $CuCl_3^{2-}$, and from $CuCl_3^{2-}$ to Cl^- can be postulated. Table 5 illustrates Markov chain evolution with a hypothetical transition probability array, indicating a slow convergence to the limiting state. The other two eigenvalues, approximately 0.06 and –0.329 have no significance, since they yield negative and larger – than - one probabilities.

4. The Electrode Surface as a Multiple Client – Server: A Markovian View

If an electrode surface is considered to contain a multiple set of identically functioning single active centres/servers, where each ion can have a choice of several adjacent servers,[28] the probability that exactly j number of severs is occupied at any given time may be computed by the Erlang formula (Eq. 15), provided that numerical values of λ and μ, or $r = \lambda/\mu$ are known. From a practical point of view, two particular states are of interest: s_0 , where the entire surface is free for an electrode reaction to proceed, and s_m , where the entire surface is covered by the reaction product; m is the number of active centres, or clusters of active centres. A small value of r represents either a slow arrival of ions, or a fast electrode reaction, and vice versa. When $r = 1$, the arrival and surface rates are matched exactly. Table 6 shows the effect of r on (i) the probability of the entire electrode surface

Table 6
Probabilities Predicted by the Erlang Formula for a Completely Free, and a Completely Occupied Surface at Various Surface Efficiencies. M = 5

$r = \lambda/\mu$	p_0^* (Eq. 23)	p_m^0 (Eq. 24)
0.1	0.905	7.54×10^{-8}
0.5	0.613	1.60×10^{-4}
1	0.368	0.00307
5	0.0109	0.284
10	6.77×10^{-4}	0.564

being free, and (ii) the probability of the entire electrode surface is being covered by a reaction product. The probabilities are computed as

$$p_0^* = 1/(\sum_{j=0}^{m} r^j / j!) \tag{23}$$

and

$$p_m^* = p_0^* r^m / m! \tag{24}$$

in the specific case of $m = 5$. For $r \leq 0.1$, $p_m^* = r^m \exp(-r)/m!$ yields a close approximation.

The estimation of cluster size is a-priori arbitrary, but it may be based on the value[29] of 10^{15} molecule/cm^2, and a representative ~ 2.25x10^{15} atom/cm^2 based on listed radii of certain metals.[30] A fraction of this quantity may be defined as an active-centre cluster. For the sake of convenient calculations, the total number of clusters should preferably be between 5 and 10.

IV. MARKOVIAN MODELING OF ELECTROLYZER (ELECTROCHEMICAL REACTOR) PERFORMANCE

1. The Batch Electrolyzer

The batch electrolyzer discussed in detail by Scott[31] for the reversible process of type $A^{z+} + ze^- \leftrightarrow B$ is employed for illustration. With no discharged product at zero time, the governing mole balances may be written as (Eqs. 2.20 and 2.21 in Ref. [31]):

$$-dA/dt = \sigma(k_{f1} - k_{b1}B)/(1 + Da_1 + Da_2) \tag{25}$$

$$dB/dt = -dA/dt \;;\; B = A_0 - A \tag{26}$$

with apparent kinetic parameters k_{f1} = 0.001 m/s, k_{b1} = 0.00091 m/s adopted from Figure 22 , p.56 in Ref. [31]. Taking k_L = 0.00005 m/s, Da_1 = 20 and Da_2 = 18.2 are the Damköhler numbers. The mole balances are rewritten as

$$10^4(dA/dt) = -1.275A + 1.160B \quad \text{mol/m}^3\text{.s}$$

$$10^4(dB/dt) = 1.275A - 1.160B \quad \text{mol/m}^3\text{.s}$$

or, in terms of fractional concentrations $u \equiv A/A_0$; $v \equiv B/A_0$; $u_0 = 1$; $v_0 = 0$,

$$10^4(du/dt) = -1.275u + 1.160v \tag{27}$$

$$10^4(dv/dt) = 1.275u - 1.160v \tag{28}$$

with analytical solution

$$u(t) = 0.4764 + 0.5236\exp(-2.436x10^{-4}t) \tag{29}$$

$$v(t) = 1 - u(t) = 0.5236[1 - \exp(-2.436x10^{-4}t)] \tag{30}$$

obtained by classical differential equation theory.

The Markovian approach is twofold. If the electrolyzer is viewed as a Markov processor, Eqs. (27) and (28) are replaced by their Kolmogorov-equivalents, Eqs. (8) and (9); the solution is obtained in terms of probabilities for state s_1 (species concentration A), and state s_2 (species concentration B). It is more inspiring (elegant?) to regard the electrolyzer as a Markov-chain generator with Markov equations

$$u(n+1) = 0.9998725u(n) + 1.160x10^{-4}v(n) \tag{31}$$

$$v(n+1) = 1.275x10^{-4}u(n) + 0.999884v(n) \tag{32}$$

Table 7
Comparison of the Markov Chain Solution (Eqs. 31 and 32) with the Analytical Solution (Eqs. 29 and 30)

Time base = 1 s		Time base = 1 h		
n	$u(n) = u(t);\ t = n$	n	$u(n)$	$u(t);\ t = n$
10	0.9987	1	0.5408	0.6943
50	0.9937	2	0.4844	0.5670
100	0.9874	3	0.4774	0.5141
200	0.9751	4	0.4766	0.4921
-	-	10	0.4765	0.4765

The analytical solution and Markov chain solution are identical for $t = n\Delta t$, provided that Δt is sufficiently small. The importance of this point is demonstrated in Table 7. At large times, discrepancy between the two solution sets gradually vanishes, i.e., when the limiting conditions are reached.

2. The Tank Electrolyzer – Mixer Cascade

In studying residence time distribution in a tank-flow electrolyzer, a tracer injected into it is recovered in a mixing tank placed downstream. If both tanks are perfectly stirred, then with a tracer pulse, the mole balance for the tracer can be written as

$$dc_1 / dt = -Qc_1 / V_1 - k_m A_e c_1 / V_1 \tag{33}$$

and

$$dc_2 / dt = Qc_1 / V_2 \tag{34}$$

where V_1 is the electrolyzer volume, V_2 the mixing tank volume, c_1 the tracer concentration in the electrolyzer, c_2 the tracer concentration in the mixing tank, and Q is the volumetric flow rate. Equation (33) implies that the tracer undergoes an electrochemical reaction with apparent mass transport coefficient k_m at an active surface A_e, under essentially mass transport-controlled conditions. The tracer is assumed to accumulate in the mixing tank (or absorbed by some manipulation), so the exit stream from the mixer does not contain tracer elements. This is a

Table 8
Comparison of the Markov Chain Solution (Eqs. 35 and 36) with the Analytical Solution (Eqs. 33 and 34). $Q/V = 2.5\text{x}10^{-3}$ min^{-1}; $(Q + k_m A_e)/V = 3.5\text{x}10^{-3}$ min^{-1}; Dimensionless Concentrations: $c_1(0) = 1$; $c_2(0) = 0$; $\Delta t = 0.1$ min; $t = 0.1n$

	Markov chain		Analytical solution[a]	
n	$c_1(n)$	$10^2\, c_2(n)$	$c_1(t)$	$10^2\, c_2(t)$
1	0.99965	0.0250	0.99965	0.0250
2	0.99939	0.0500	0.99930	0.0500
5	0.99825	0.1249	0.99825	0.1249
10	0.99651	0.2500	0.99651	0.2495
50	0.98265	1.2390	0.98265	1.2390
100	0.96560	2.4570	0.96560	2.4560
200	0.93238	4.8300	0.93239	4.8280
500	0.89943	11.469	0.83946	11.466

[a] $c_1(t) = \exp(-3.5\text{x}10^{-3}t)$
$c_2(t) = 0.7142\exp[1 - \exp(-3.5\text{x}10^{-3}t)]$
$[t]$: min

modification of the configuration discussed by Tamir[1] (Section 4.3-1, pp. 353–355) via the inclusion of the electrochemical process. Rewriting the mole balances as $dc_1 / dt = -\mu_1 c_1; \mu_1 \equiv \equiv (Q + k_m A_e)/V_1$ and $dc_2 / dt = \mu_2 c_1; \mu_2 \equiv Q/V_2$, the Markov equations are obtained as

$$c_1(n+1) = (1 - \mu_1 \Delta t) c_1(n) \tag{35}$$

$$c_2(n+1) = (\mu_2 \Delta t) c_1(n) + c_2(n) \tag{36}$$

If $\Delta t = 1$ in terms of a convenient time unit, $c_1(n+1) = (1 - \mu_1)c_1(n)$, and $c_2(n + 1) = \mu_2 c_1(n) + c_2(n)$. The limiting conditions yield $c_1^* = 0$ and $c_2^* = 1$, if dimensionless concentrations are employed, using e.g., the initial amount of tracer concentration as reference. Table 8 indicates very good agreement up to large times between the analytical and Markov chain solution, when the size of Δt is judiciously chosen.

3. Markovian Analysis of the CSTER[23]

In the Markovian view of the CSTER (continuous-flow stirred tank electrochemical reactor) transitional probabilities related to flow and reaction describe the history of an element (molecule, ion, particle, etc.). A convenient framework for analysis considers the interior of the reactor as state s_1, and exit as s_2, as an absorbing state. Then, $p_{11}(n)$ denotes the probability that an element has not undergone reaction, and is inside the reactor at stage n (i.e., time instant $n\Delta t$). Consequently, $p_{12}(n)$ is the probability that the element at stage n has left the reactor, while it had been in the reactor at stage $n - 1$. The probability that at stage n the element was in the reactor, while it had been in the exit stream at stage n-1, is $p_{21}(n)$. Finally, $p_{22}(n)$ denotes the probability that an element in the exit stream a stage n was in the exit stream at stage $n - 1$. In the absence of backmixing or recycling, $p_{21}(n) = 0$. In a perfectly mixed reactor the transitional probability of outflow over a time period Δt is $Q\Delta t/V$, if Q is the volumetric flow rate and V is the active volume.

In this context, the mole balance for a reacting element (e.g., active ion):

$$dc_1 / dt = Qc_i / V - Qc_1 / V - k_m A_e c_1 / V \tag{37}$$

may be rewritten as a Markov equation

$$c_1(n+1) = x_i(n+1) + p_{11}c_1(n) \tag{38}$$

where $x_i(n) = Q\Delta t c_i(n)/V$ represents the reacting element inflow with inlet concentration $c_i(t)$, and $p_{11} = 1 - [(Q + k_m A_e)/V]\Delta t$. If the inlet concentration remains constant, $x_i(n) = x_i$ does not vary with n.

The importance of choosing a suitably small Δt is again shown in Table 9; if Δt is too large, the Markovian approach yields a meaningless temporary rise in c_1. Nonetheless, the limiting value $c_1^* = 9.663\text{x}10^{-3}$ mol dm^{-3} is predicted correctly.

In a more sophisticated version, the analysis is based on three states. The transition probability p_{13} pertains to the bypass of an element entering the reactor to the exit state without participating in the electrode process. To analyze a CSTER with partial

Table 9
Illustration of the Importance of Suitably Small Δt Values for Markovian Analysis of a CSTER.
$Q = 2\ dm^3\ min^{-1}$;
$c_i = 0.01\ mol\ dm^{-3}$; $V = 10\ dm^3$;
$A_e = 700\ cm^2$; $k_m = 10^{-4}\ cm\ s^{-1}$;
$c_1(0) = 0.01\ mol\ dm^{-3}$; $t = n\Delta t$

	Δt = 1 min		Δt = 0.1 min
n	$c_1(n)$	$c_2(n)$	$c_1(t) = c_2(t)$
0	10	10	10
1	12	9.94	9.99
2	11.52	9.88	9.99
3	29.13	9.84	9.98
4	25.10	9.81	9.97
5	21.91	9.78	9.97

All concentrations in mmol dm^{-3} units
$c_1^* = 9.66$

backmixing,a four-state model has been shown to be attractive (Section 5, Ref. [23]).

V. THE MARKOV–CHAIN APPROACH TO MISCELLANEOUS ASPECTS OF ELECTROCHEMICAL TECHNOLOGY

Markov theory can be applied advantageously to a variety of technological problems of day-to-day operation in an industrial environment. Such problems transcend the realm of purely physico-chemical and fundamental electrochemical considerations, and are closely linked to (large-scale) technological aspects. This Section demonstrates the utility of the Markovian approach to selected examples.

1. Repair of Failed Cells

An electrolytic cell bank producing a commercial product consists of m number of cell units. Due to various reasons (e.g., imperfect

control of electrolyte composition and/or purity), cells can be short-circuited temporarily. It is assumed that the Poissonian failure-density parameter is λ, i.e., that the probability of short-circuiting of a cell during time interval Δt is $\lambda \Delta t$. The repair of short-circuited cells requires time whose distribution is exponential with density parameter μ.

In the simplest case (suppressing the trivial situation of no repair being available), there is one repair *station* (mechanic, technician, etc.) available at any given time. The task is to determine the limiting probability distribution of j cells being out of order, assuming time-independent conditions. The starting point of this procedure is Eq. (13), stated again as

$$\sum_{i=0}^{m} p_i^* \lambda_{ij} = 0; j = 0,...,m \tag{13}$$

The slightly changed notation reflects the choice of states: s_0 is the state of normal operation, and state s_j is the state of being out of order due to short-circuiting of j number of cells. Thus, λ_{00} represents normal operation, and λ_{0j} represents the breakdown of j number of cells, etc. Δt is sufficiently small for only one cell to become short-circuited, or repaired within this time interval. For the $j = 0$ case, Eq. (13) expands to

$$p_0^* \lambda_{00} + p_1^* \lambda_{10} = 0 \tag{39a}$$

and since $\lambda_{00} = -m\lambda$, $\lambda_{10} = \mu$, it follows that

$$m\lambda p_0^* = \mu p_1^* \tag{39b}$$

If $m = 1$, Eq. (13) yields $p_0^* \lambda_{01} + p_1^* \lambda_{11} = 0$. Since $\lambda_{01} = \lambda$ and $\lambda_{11} = \mu$, $\lambda p_0^* = \mu p_1^*$. This relationship can more simply be obtained by setting $m = 1$ in Eq. (39b). If $m = 2$, Eq. (13) expands to:

$$j = 1: \quad p_0^* \lambda_{01} + p_1^* \lambda_{11} + p_2^* \lambda_{21} = 0 \tag{40a}$$

$$j = 2: \quad p_0^* \lambda_{02} + p_1^* \lambda_{12} + p_2^* \lambda_{22} = 0 \tag{40b}$$

Since $\lambda_{01} = 2\lambda$; $\lambda_{02} = 0$; $\lambda_{11} = -(\lambda + \mu)$; $\lambda_{12} = \lambda$; $\lambda_{21} = \mu$; $\lambda_{22} = -\mu$; the final forms become

$$j = 1: \quad (\lambda + \mu) p_1^* = 2\lambda p_0^* + \mu p_2^* \tag{41a}$$

$$j = 2: \quad \lambda p_1^* = \mu p_2^* \tag{41b}$$

Table 10 summarizes the reasoning behind the establishment of the λ_{ij} values. By successive development, the expansion set

$$j = 1: \quad (2\lambda + \mu) p_1^* = 3\lambda p_0^* + \mu p_2^* \tag{42a}$$

$$j = 2: \quad (\lambda + \mu) p_2^* = 2\lambda p_1^* + \mu p_3^* \tag{42b}$$

$$j = 3: \quad \lambda p_2^* = \mu p_3^* \tag{42c}$$

is obtained for the case of $m = 3$. In the general case, the limiting probability distribution is given by the set of equations

$$m\lambda p_0^* = \mu p_1^* \tag{43a}$$

$$[(m - j)\lambda + \mu] p_j^* = (m - j + 1)\lambda p_{j-1}^* + \mu p_{j+1}^* \tag{43b}$$

Table 10
Development[32] of the Transitional Probability Density Parameters of Eq. (13); $m = 2$

Relationship	Substantiation
$\lambda_{01} = 2\lambda$	Either one of the two cells can break down
$\lambda_{02} = 0$	Only one cell can break down within Δt
$\lambda_{11} = -(\lambda + \mu)$	No new breakdown; no new repair
$\lambda_{12} = \lambda$	Second cell breaks down, first cell not repaired
$\lambda_{21} = \mu$	One cell repaired; one cell still out of order
$\lambda_{22} = -\mu$	Neither cell has been repaired

$$\mu p_m^* = \lambda p_{m-1}^* \tag{43c}$$

Equations (43) are also stated by Rozanov[5] (p. 113, Problem 6) without derivation.

The limiting probability of exactly j cells being out of order is obtained by setting $j = 0, 1,\ldots, m$ in Eq. (43) and recalling that all probabilities add up to unity, as[5]

$$p_j^* = [m!/(m-j)!](\lambda/\mu)^j p_0^*; j = 0,1,...,m \tag{44}$$

and the mean number of cells requiring repair[5] (Problem 7, p.113; no derivation) is

$$n_m = m - [(\lambda/\mu + 1)/(\lambda/\mu)](1 - p_0^*) \tag{45}$$

A typical distribution is portrayed in Table 11 for a configuration of fifty cells arranged in five ten-cell banks, where for the sake of safety, a short-circuited cell will disconnect an entire bank. This arrangement is essentially identical to a five-cell bank with single short-circuits. The more efficient (i.e., the faster) the repair service (i.e., the smaller λ/μ), the smaller the probability that short-circuited cells are not repaired at any arbitrary time. Conversely, if repair takes a long time, a certain number of cells will always be short-circuited, and if λ/μ is extremely large, a complete shutdown of cell operation will be highly probable (the $\lambda/\mu = 10$, $p_5^* \approx 0.9$ case is a case in point, albeit not overly realistic.)

Table 11
Limiting Probability Distribution for Exactly j = 0,1,…,5 Cells being Short-Circuited, as a Function of Repair Efficiency λ/μ

	p_j						n_m
λ/μ	$j=0$	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	
0.1	0.564	0.282	0.113	0.034	6.8×10^{-3}	6.8×10^{-4}	0
0.5	0.0367	0.0917	0.183	0.275	0.275	0.138	2.11
1	3.7×10^{-3}	0.0153	0.0614	0.184	0.368	0.368	3.01
5	2.2×10^{-6}	5.46×10^{-5}	1.09×10^{-3}	0.0164	0.164	0.819	3.8
10	7.5×10^{-8}	3.77×10^{-7}	1.51×10^{-4}	4.53×10^{-3}	0.090	0.905	3.9

2. Analysis of Switching–Circuit Operations

An electrolytic plant (e.g., a metal refinery) may operate m number of cells which have to be taken out of service temporarily at regular time intervals. The probability that a cell is being put back into service within a small Δt time period is $\lambda\Delta t$, where λ is the density parameter. The probability that a cell is taken out of operation within Δt is $\mu\Delta t$, and μ is the second density parameter. The limiting probabilities p_j^* ; $j = 0,1,2\ldots,m$ of exactly j number of cells being switched back into operation can be obtained by a procedure identical to the electrical power line/machine operation problem stated by Rozanov[5] (Problem 9, p. 114) and Sveshnikov[6] (Problem 39.16, pp. 255–256; 455):

$$m\lambda p_0^* = \mu p_1^* \tag{46a}$$

$$[(m-j)\lambda + j\mu]p_j^* = (m-j+\lambda)p_{j-1}^* + (j+1)\mu p_{j+1}^* \tag{46b}$$

$$m\mu p_m^* = \lambda p_{m-1}^* \tag{46c}$$

whose derivation is similar to the one in Section V.1 . Systematic iteration yields

$$p_j^* = \{m!/[(j!(m-j)!]\}[\mu/(\lambda+\mu)]^{m-j}[\lambda/(\lambda+\mu)]^j \quad ; \; j = 0,1,2,\ldots,m \tag{47}$$

Table 12 illustrates the computation procedure in the case of $m = 5$; the plant may be envisaged, as in Section V.1, to consist of m cell banks, the quantity j denoting the number of banks switched back into operation. In the specific case of $\lambda = \mu$ (equi - probability of switching into either direction), Eq. (47) reduces to the binomial probability distribution of selecting j elements out of m identical elements with a single-event probability of ½.

Table 12
Limiting Probability Distribution for Exactly $j = 1, 2,\ldots, 5$ Cells being Switched Back into Operation, as a Function of Switch Parameter, λ/μ

λ/μ	p_j					
	$j = 0$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$
0.1	0.621	0.310	0.062	$6.21\text{x}10^{-3}$	$3.10\text{x}10^{-4}$	$6.20\text{x}10^{-6}$
0.5	0.132	0.329	0.329	0.165	0.041	$4.10\text{x}10^{-3}$
1	0.031	0.156	0.312	0.312	0.157	0.031
5	$1.29\text{x}10^{-4}$	0.0032	0.032	0.161	0.402	0.402

3. Reliability of Warning Devices

An electrochemical plant is equipped with m number of warning devices to signal component failure in electrolytic processes. The distribution of failure probability is Poisson with density parameter λ, while the signal time of the warning devices has an exponential distribution with density parameter μ. The devices are identical and operate independently of one another. To determine the limiting probability that a component failure will *not* be signaled it is advantageous to employ the Erlang formula (Eq. 15) written for the m devices:

$$p_m^* = (1/m!)(\lambda/\mu)^m / [\sum_{j=0}^{m} (1/j!)(\lambda/\mu)^j] \tag{48}$$

Table 13
The Effect of the λ/μ Efficiency Parameter on the Probability of Signaling Failure as a Function of the Number of Warning Devices

λ/μ	p_m^*				
	$m = 2$	$m = 3$	$m = 4$	$m = 5$	m = 6
½	0.143	0.023	$2.9\text{x}10^{-3}$	$3.6\text{x}10^{-5}$	$3.0\text{x}10^{-6}$
1	0.200	0.045	0.015	$3.1\text{x}10^{-3}$	$5.1\text{x}10^{-4}$
2	0.400	0.210	0.095	0.037	0.012

to ascertain the number of warning devices needed in order to ensure that the probability of failure-signaling does not exceed a predetermined value. As shown in Table 13, p_m^* is highly sensitive to the device efficiency (λ/μ).

4. Monitoring of Parasitic Reaction Sequence in an Electrolyzer

In a complex reaction system occurring in a flow electrolyzer, one of the anode reaction products reacts with the principal constituent of the composite electrolyte, yielding an anionic species whose parasitic reaction at the cathode reduces the cathode efficiency.

Due to system complexity and paucity of information about the reaction mechanism(s), a Markov process model is proposed with probability $p_k = \lambda^k \exp(-\lambda)/k!$ of the production of k ions of the anionic species; λ is the mean of the postulated Poisson distribution. The model also stipulates a probability q of reaction of the anionic species during two successive equal time periods of Δt each.

The limiting probabilities p_j^* for the number of ionic species in the system can be established in a manner similar to the production/decomposition of a radioactive particle possessing the same probabilistic properties[6] (Example 38.3, pp. 237–238). The result of an involved algebraic manipulation, itself a Poisson distribution:

$$p_j^* = (\mu^j / j!)\exp(-\mu); \mu \equiv \lambda / q; j = 0,1,... \tag{49}$$

is not explicitly dependent on k. As indicated previously, the number of ionic species can conveniently be interpreted to mean a cluster of ions, whose (arbitrary) size depends on the properties of the electrochemical system examined. Table 14 depicts the rapid increase in the number of (unreacted) anions with an increase in μ. When, e.g., $\mu = 10$, $p_9^* = p_{10}^* = 0.125$ is the largest individual probability, but the descent of p_j^* from the maximum value is slow [$p^*(15) = 0.035$; $p(20)^* = 0.0019$; $p(25)^* = 2.9\text{x}10^{-5}$.] Cumulative probabilities are expectably higher, e.g., the probability that, when

Table 14
Limiting Probability Distribution for Exactly $j = 0, 1, 2,\ldots$ Number of an Ionic Species Existing in the Electrolyte, as a Function of the λ/q Ratio in Eq. (49)

μ/λ	p_j^*						
	$j = 0$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
0.1	0.905	0.091	$4.5\text{x}10^{-3}$	$2.0\text{x}10^{-4}$	$3.8\text{x}10^{-6}$	$7.5\text{x}10^{-8}$	$2.5\text{x}10^{-17}$
0.5	0.607	0.303	0.076	0.013	$1.6\text{x}10^{-3}$	$2.0\text{x}10^{-4}$	$1.6\text{x}10^{-10}$
1.0	0.368	0.368	0.184	0.061	0.015	$3.1\text{x}10^{-3}$	$1.0\text{x}10^{-7}$
5.0	$6.7\text{x}10^{-3}$	0.034	0.084	0.140	0.176	0.176	0.018
10.0	$4.5\text{x}10^{-5}$	$4.5\text{x}10^{-4}$	$2.3\text{x}10^{-3}$	$7.6\text{x}10^{-3}$	0.019	0.038	0.125

$\mu = 0.1$, not more than three anions/anion clusters are present is virtually 100%.

5. Flow Circulation in an Electrolytic Cell

The interaction of forced and natural convective flow between cathodes and anodes may produce unusual circulation patterns whose description via deterministic flow equations may prove to be rather unwieldy, if possible at all. The Markovian approach would approximate the true flow pattern by subdividing the flow volume into several zones, and characterize flow in terms of transition probabilities from one zone to others. Under steady operating conditions, they are independent of stage n, and the evolution pattern is determined by the initial probability distribution. In a similar fashion, the travel of solid pieces of impurity in the cell can be monitored, provided that the size, shape and density of the solids allow the pieces to be *swept* freely by electrolyte flow.

Table 15 illustrates the development of the probability distribution of a solid particle being in one of the three characteristic zones in a cell, assuming that the particle residence has an equiprobable initial state, and that the transitional probabilities are as follows. Zone 1 →2:0.4; Zone 2 →1:0.4; Zone 1→3: 0.1;Zone 3→1:0.2; Zone 2→3:0.3; Zone 3→2:0.2 .

The continuous circulation of flow elements inside the electrolyzer is due to the donor-acceptor nature of each zone. If, however, at least one of the zones is an absorbing Markov state,

Table 15
Probability Distribution of a Solid Particle being in One of Three Flow Circulation Zones with an Equiprobable Initial-State Probability.
$p_{11} = 0.5$; $p_{12} = 0.4$; $p_{13} = 0.2$; $p_{21} = 0.4$; $p_{22} = 0.3$; $p_{23} = 0.6$; $p_{31} = 0.1$; $p_{32} = 0.3$; $p_{33} = 0.2$

n	$p_1(n)$	$p_2(n)$	$p_3(n)$
0	1/3	1/3	1/3
1	0.367	0.433	0.200
2	0.397	0.397	0.206
3	0.398	0.402	0.200
4	0.400	0.400	0.200
∞	0.400	0.400	0.200

the distribution pattern eventually excludes all states except the absorbing state(s). If the exit zone is the only absorbing state, the probability that flow elements/solid particles will eventually flow out of the electrolyzer is 100%. Tables 16 and 17 indicate the slow rate whereby the limiting probabilities are reached. The initial state probability distribution is immaterial, since the eigenvector associated with eigenvalue of unity (Section II.1) has one element of unity, irrespective of the p_{ij} values (Appendix), while the other elements are zero.

VI. FINAL REMARKS

Unidirectional kinetic processes cannot be immediately interpreted as Markov chains, since only the (1,1) element of the *P*-matrix would differ from zero, violating the stochastic matrix constraints (Section II.1). An artificial Markov matrix complying with this constraint can be visualized, however, with the understanding that no other element of this *imbedded P*-matrix, past the (1,1) element, will have a physical meaning. It follows that the initial state probability vector is non-zero only in its (1,1)

Table 16
Probability Distribution of a Solid Particle being in One of Three Flow Circulation Zones with an Initial Distribution of $p_1^0 = 1$; $p_2^0 = 0$; $p_3^0 = 0$. The Transition Probabilities are: Zone 1→2: 0.2; Zone 1→3: 0.1; Zone 2→3: 0.05

n	$p_1(n)$	$p_2(n)$	$p_3(n)$
0	0.700	0.200	0.100
1	0.700	0.200	0.100
5	0.168	0.485	0.347
10	0.028	0.456	0.515
20	$8.0\text{x}10^{-4}$	0.286	0.713
30	$2.2\text{x}10^{-5}$	0.172	0.828
50	$1.8\text{x}10^{-8}$	0.062	0.938
100	$3.2\text{x}10^{-16}$	0.005	0.995
150	$5.8\text{x}10^{-24}$	$3.6\text{x}10^{-24}$	0.9996

element, and that it has to be unity. The small-signal analysis[36] and large-step coulostatic analysis[37] employed in electroanalytical chemistry provide a good case in point.

The variation of overpotential with time is given by the expression

Table 17
Probability Distribution in Table 16 Modified by Additional Transitional Probability for Zone 2→ 1: 0.1

n	$p_1(n)$	$p_2(n)$	$p_3(n)$
0	0.700	0.200	0.100
1	0.700	0.200	0.100
5	0.255	0.392	0.353
10	0.141	0.314	0.545
20	0.069	0.163	0.768
30	0.035	0.083	0.881
50	0.009	0.022	0.969
100	$3.23\text{x}10^{-4}$	$7.59\text{x}10^{-4}$	0.999
150	$1.13\text{x}10^{-5}$	$2.65\text{x}10^{-5}$	0.99996
∞	0	0	1

$$d\eta / dt = -(1/\tau_c)\eta; \eta(t=0) = \eta_0 \tag{50}$$

where the time constant $\tau_c = RTC_d/nFi_0$ is a lumped parameter incorporating the surface capacity C_d, and exchange current density i_0. The Markovian approach assigns an imbedded transition probability matrix with elements $p_{11} = 1 - \Delta t/\tau_c$, $p_{12} = 0$; $p_{21} = \Delta t/\tau_c$, $p_{22} = 1$, and the initial state probability vector has elements $p_1^0 = 1$; $p_2^0 = 0$. Consequently, the stage-dependent elements are $p_{11}(n) = (1 - \Delta t/\tau_c)^n$, and $p_{12}(n) = 1 - (1 - \Delta t/\tau_c)^n$, the latter being irrelevant. Comparison with the analytical solution of Eq. (50) in Table 18 indicates, as expected, that close agreement can be attained if Δt is sufficiently small.

It is instructive to note that the standard first-order forward difference discretization of Eq. (50) is finite difference calculus:

$$d\eta / dt \cong [\eta(t + \Delta t) - \eta(t)] / \Delta t \tag{51}$$

also leads to the Markov chain solution in the second column of Table 18.

Table 18
Comparison of the Analytical Solution of Eq. (50) and the Markov Chain Solution.
T = 293 K; C_d = 2x10^{-5} F cm^{-2}; i_0 = 10^{-5} A cm^{-2}; n = 2; R = 8.3143 J mol^{-1} K^{-1}; F = 96488 C mol^{-1}; Δt = 0.1 ms; τ_c = 25.25 ms

n	Markov chain: $Y(n) = (0.99604)^n$	Analytical solution:[a] $Y(0.1n) = \exp(-0.00396n)$
0	1	1
1	0.99604	0.99605
2	0.99210	0.99211
3	0.98817	0.98819
4	0.98425	0.98428
5	0.98036	0.98039
10	0.96110	0.96117
50	0.82005	0.82037
100	0.67248	0.67301
1000	0.01891	0.01906

[a] $Y = \eta / \eta(0)$

At this time, it would be hazardous to predict the eventual extent of influence Markov theory will attain in the realm of electrochemistry and electrochemical engineering. Much terrain remains to be explored, especially in the potential of multidimensional chains and processes. If this chapter succeeds in kindling at least some sustained interest in the subject matter, it will have reached its ultimate objective.

ACKNOWLEDGMENTS

Author's work in various domains of probability theory applied to electrochemical phenomena has been supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), and the University of Waterloo.

APPENDIX

Eigenvalues and Eigenvectors of the Transition Probability Matrix

Under limiting conditions, the state probabilities become invariant, i.e., past a sufficiently large value $n > N$, $\boldsymbol{p}(n+1) = \boldsymbol{p}(n)$; $n = N, N+1; N+2,\ldots$, hence $\boldsymbol{p}(n) = \boldsymbol{P}\boldsymbol{p}(n)$. Invoking the classical eigenvalue theorem of matrix algebra $\lambda\boldsymbol{e} = \boldsymbol{P}\boldsymbol{e}$ for the Markov matrix $\boldsymbol{P}$ results in the important finding that that the limiting probabilities are elements of the eigenvector $\boldsymbol{e}$ pertaining to the specific eigenvalue $\lambda = 1$. However, the unity eigenvalue is not the only one, because the total number of eigenvalues equals the number of states. The $\lambda = 1$ is, however, the only admissible eigenvalue of **P**.

As an illustration, the oscillatory behaviour of anodic copper dissolution discussed in Section III.3, with $\boldsymbol{P}$ elements shown in Table 5, is chosen. The eigenvalue theorem yields three relationships ($k = 1,2,3$):

$$0.1e_{1k} + 0.5e_{2k} + 0.4e_{3k} = \lambda_k e_{1k} \tag{52}$$

$$0.8e_{1k} + 0.1e_{2k} = \lambda_k e_{2k} \tag{53}$$

$$0.1e_{1k} + 0.4e_{2k} + 0.6e_{3k} = \lambda_k e_{3k} \tag{54}$$

To find p_k^*, k = 1,2,3 λ = 1 is set and the solution of the above equations yields p_1^* = 36/109 ≈ 0.3303; p_2^* = 32/109 ≈ 0.2936; p_3^* = 41/109 ≈ 0.3761.

The eigenvalues are given by the characteristic equation of the ***P*** matrix:

$$\lambda^3 - 0.8\lambda^2 - 0.31\lambda + 0.11 = 0 \tag{55}$$

as λ_1 = 1; $\lambda_2 \approx 0.0567$ and $\lambda_3 \approx -0.3285$. The eigenvectors for the second and third eigenvalues are, respectively: {–0.2311; 4.2695; –3.1008}, and {–5.7669; 10.784; –4.0172}. Since probabilities cannot be negative, and cannot be larger than unity, these eigenvector elements cannot be interpreted as probabilities. The limiting probabilities are the elements of the first eigenvector {36/109; 32/109; 41/109}.

The eigenvalue theorem, a standard topic in any linear algebra text and mathematics book, is particularly well presented by Dixon,[33] Wilf,[34] and Nicholson.[35]

REFERENCES

[1]A. Tamir, *Application of Markov Chains in Chemical Engineering*, Elsevier, Amsterdam, 1998.
[2]V. S. Pugachev, *Teoriya Sluchainikh Funktzii* (Theory of Random Functions), Gos. Izd. Fiziko-Matem. Liter., 2nd Ed.., Moscow, 1960, Section 46, pp. 191–202.
[3]E. Parzen, *Modern Probability Theory and Its Applications*, Wiley and Sons, New York, 1960, Chapter 3, Section 3, pp. 136–147.
[4]H. Stark and J. W. Woods, *Probability, Random Processes and Estimation Theory for Engineers*, Prentice Hall, Englewood Cliffs, 1986.
[5]Y. A. Rozanov, *Probability Theory: A Concise Course*, Dover, New York, 1969, Chapter 7, pp. 83–101; Chapter 8, pp.102 – 114.
[6]A. A. Sveshnikov, *Problems in Probability Theory, Mathematical Statistics and Theory of Random Functions*, Dover, 1968, Chapter VIII, pp. 231–274.
[7]F. R. Gantmacher, *The Theory of Matrices*, Chelsea Publishing, New York, 1959, Vol. 2, Chapter XIII, Sections 6 – 8, pp. 80 – 98.
[8]M . D. Greenberg, *Advanced Engineering Mathematics*, 2nd Ed., Prentice Hall, Upper Saddle River, 1998, Section 11.2, pp. 546 – 549.

[9]W. J. Stewart, *Introduction to the Numerical Solution of Markov Chains*, Princeton Univ. Press, 1994.
[10]A. Sinclair, *Algorithms for Random Generation and Counting*: *a Markov Chain Approach*, Birehäuser, Boston, 1993.
[11]W. J. Anderson, *Continuous-Time Markov Chains: An Applications-Oriented Approach*, Springer–Verlag, New York, 1991.
[12]O. Hernandez-Lerma, *Adaptive Markov Control Processes*, Springer-Verlag, New York, 1989.
[13]G. Winkler, *Image Analysis*, *Random Fields and Markov Chain Monte Carlo Methods*, Springer-Verlag, New York, 2003.
[14]P. Baldi, L. Mazliak, and P. Priouret, *Martingales and Markov Chains*, Chapman and Hall/CRC, Boca Raton, 2002.
[15]T. Z. Fahidy, *Can. J. Chem. Eng*. **65** (1987) 1009.
[16]G. Radons, J. D. Becker, B. Duelfer, and J. Krueger, *Biol. Cybern*. **71** (1994) 359.
[17]A. Oprişan, S.-A. Oprişan, and M. Ignat, *Analele Stiinţifice Ale Universităţii "Al.I. Guza"*, (in English), Iasi (Romania), 1997 – 1998.
[18]S. A. Godorr, B. D. Young and A. W. Bryson, *Chem. Eng. Commun*. **117** (1992) 307.
[19]T. Witten and L. M. Sauder, *Phys. Rev. Lett*. **47** (1981) 1400.
[20]T. Witten, and L. M. Sauder, *Phys. Rev. B* **27** (1983) 5686.
[21]F. Argoul, A. Arneodo, G. Grasseau, and H. L. Swinney, *Phys. Rev. Lett*. **61** (1988) 2558.
[22]M. R. Dudek, W. M. Trochimczuk, and R. Wycizk, *J. Membrane Sci*. **67** (1992) 273.
[23]T. Z. Fahidy, *J. Appl. Electrochem*. **17** (1987) 841.
[24]L. J. Goldstein, and D. I. Schneider, *Finite Mathematics and Its Applications*, 2nd Ed., Prentice Hall, Englewood Cliffs, 1984, Chapter 8, pp. 329 – 360.
[25]T. Z. Fahidy, *Chem. Eng. Edu*. **27** (1993) 42.
[26]D. C. Montgomery, G. C. Runger, and N. F. Hubele, *Engineering Statistics*, 2nd Ed., Wiley and Sons, New York, 2001, Section 2.1, Exercise 2.3, p. 25.
[27]Z. H. Gu, J. Chen, and T. Z. Fahidy,*Electrochim. Acta* **37** (1992) 2637.
[28]T. Z. Fahidy, *J. Appl. Electrochem*. **28** (1998) 411.
[29]J. O'M. Bockris, A. K. N. Reddy, and M. Gamboa-Aldeco, *Modern Electrochemistry*, 2nd Ed., Plenum, New York, Volume 2A: Fundamentals of Electrodics, Section 6.2.6.2, p. 804.
[30]L. H. Van Vleck, *Elements of Materials Science and Engineering*, 4th Ed., Addison-Wesley, Reading, Mass., 1980, Appendix B, pp. 520 -523.
[31]K. Scott, *Electrochemical Reaction Engineering*, Acad. Press, London, 1991, Section 2.1.2, pp. 52 – 56.
[32]T. Z. Fahidy, unpublished results
[33]C. Dixon, *Linear Algebra*, Van Nostrand Reinhold, New York, 1971, Section VII.1, pp. 135 – 140.
[34]H. S. Wilf, *Mathematics for the Physical Sciences*, Dover, New York, 1962, Section 1,12, pp. 18 – 20.
[35]W. K. Nicholson, *Elementary Linear Algebra With Applications*, Prindle, Weber and Smith, Boston, 1986, Section 7.1, pp. 311 – 324
[36]A. J. Bard, and L. R. Faulkner, *Electrochemical Methods*, Wiley and Sons, New York, 1980, Section 7.7.2, pp. 273 – 274.
[37]Idem, ibid. Section 7.7.3, pp. 275 – 276.

4

Fractal Approach to Rough Surfaces and Interfaces in Electrochemistry

Joo-Young Go and Su-Il Pyun*

Department of Materials Science and Engineering, Korea Advanced Institute of Science and Technology, 373-1, Guseong-dong, Yuseong-gu, Daejeon, 305-701, Republic of Korea
**Corresponding author; Tel.: +82-42-869-3319; Fax: +82-42-869-3310; E-mail: sipyun@webmail.kaist.ac.kr*

I. INTRODUCTION

The real objects found in nature have complex structures which Euclidean geometry can not characterize. After Mandelbrot[1] developed a new geometry, i.e., fractal geometry, which provides a new paradigm for understanding many physical phenomena in nature, fractal geometry has been widely used in a number of fields, e.g., science, art,[2-4] economics,[5-8] etc. Especially, in science, the characterization of the rough surfaces and interfaces using fractal geometry has played an important role in understanding the anomalous behavior of rough surfaces and interfaces.

There are three kinds of fractals as shown in Figure 1.[9-11] The first one is a surface fractal, i.e., a dense object with a fractal

Modern Aspects of Electrochemistry, Number 39, edited by C. Vayenas et al., Springer, New York, 2005.

Figure 1. Schematic diagrams of three kinds of fractals. (a) A surface fractal, (b) a mass fractal, and (c) a pore fractal.

surface (Figure 1a). The second one is a mass fractal, i.e., a fractal object such as a network or cluster, whose surface is also fractal (Figure 1b). Finally, the last one is a pore fractal, i.e., a dense object in which there exists a distribution of holes or pores with a fractal structure (Figure 1c). Since it is one of the most interesting issues to estimate accurately real active area in an electrochemical system, many researchers have given a lot of attention to surface fractal as a powerful tool to quantify the geometrical irregularity of the electrode surfaces and various interfaces. In this respect, we focus on surface fractal rather than mass and pore fractals in this article.

Diffusion-limited electrochemical techniques as well as physical techniques have been effectively used to determine the surface fractal dimensions of the rough surfaces and interfaces made by electrodeposition,[12-26] fracture,[27-32] vapor deposition,[21,33-46] corrosion,[47-54] surface modification such as machining and plasma irradiation,[55-60] etc. Many rough surfaces and interfaces are self-similar fractals which show isotropic scaling behavior in all directions, but most rough surfaces and interfaces are more complex self-affine fractals which show asymmetric scaling behavior perpendicular to the surface.

Since self-similar and self-affine fractals show different scaling properties, the surface fractal dimension of the rough surfaces and interfaces must be determined based on their inherent scaling properties. However, a little attention has been paid to clarify whether the rough surfaces and interfaces show self-similar scaling property or self-affine scaling property and whether the physical and electrochemical methods for the determination of the surface fractal dimension are adequately used or not. In this respect, we review fractal approaches to rough surfaces and interfaces in electrochemistry to give clear guidelines for their characterization using the surface fractal dimension.

The structure of this review is composed of as follows: in Section II, the scaling properties and the dimensions of self-similar and self-affine fractals are briefly summarized. The physical and electrochemical methods required for the determination of the surface fractal dimension of rough surfaces and interfaces are introduced and we discuss the kind of scaling property the resulting fractal dimension represents in Section III.

In Section IV, from the studies on diffusion towards self-affine fractal interface, the surface fractal dimension as determined by the electrochemical method is characterized as being self-similar, even though the rough surfaces and interfaces show the self-affine scaling property. Finally, in Section V, we exemplified the application of fractal geometry in electrochemical systems in view of the characterization of rough surfaces and interfaces by the surface fractal dimension.

II. OVERVIEW OF FRACTAL GEOMETRY

The word 'fractal' was coined by Mandelbrot in his fundamental book.[1] It is from the Latin adjective *fractus* which means broken and it is used to describe objects that are too irregular to fit into a traditional geometrical setting. The most representative property of fractal is its invariant shape under self-similar or self-affine scaling. In other words, fractal is a shape made of parts similar to the whole in some way.[61] If the objects are invariant under isotropic scale transformations, they are self-similar fractals. In contrast, the real objects in nature are generally invariant under anisotropic transformations. In this case, they are self-affine fractals. Self-affine fractals have a broader sense than self-similar fractals. The distinction between the self-similarity and the self-affinity is important to characterize the real surface in terms of the surface fractal dimension.

From another point of view, fractals fall naturally into two categories, i.e., nonrandom and random.[1,62-64] When the enlarged part overlaps the original object exactly, we call this object a nonrandom fractal, viz., deterministic fractal. For example, the von Koch curve, the Cantor set, the Sierpinski gasket, etc. which are generated by iterative operation according to a growth rule belong to nonrandom fractals. Their self-similar or self-affine scaling property is satisfied over all range of length scales.

However, nonrandom fractals are not found in nature. Many objects existing in nature belong to random fractal, viz., statistical fractal, when the enlarged part is similar to the original object in a statistical sense. In practice, it is impossible to verify that all moments of the distributions are identical, and claims of statistical

self-similarity or self-affinity are usually based upon only a few moments, i.e., the self-similar or self-affine scaling property of random fractal is satisfied over a significant range of length scales. The classical examples are coastlines, clouds, mountains, etc.

Now we will introduce briefly the concept of self-similar and self-affine fractals by considering the assumption that fractals are sets of points embedded in Euclidean E-dimensional space.

1. Self-Similar Fractal[61,62]

The similarity transformation transforms a set of points S at position $\mathbf{x} = (x_1, ..., x_E)$ in Euclidean E-dimensional space into a new set of points $r(S)$ at position $\mathbf{x}' = (rx_1, ..., rx_E)$ with the same value of the scaling ratio $0 < r < 1$. A bounded set S is self-similar with respect to a scaling ratio r if S is the union of N non-overlapping subsets $S_1, ..., S_N$, each of which is congruent to the set $r(S)$. Here congruent means that the set of points S_i is identical to the set of points $r(S)$ after possible translations and/or rotations. For the deterministic self-similar fractal, the self-similar fractal dimension $d_{F,ss}$ is clearly defined by the similarity dimension $d_{F,S}$ given by

$$d_{F,S} = \frac{\ln N}{\ln 1/r} \tag{1}$$

Figure 2a shows a deterministic self-similar fractal, *i.e.*, the triadic Koch curve, generated by the similarity transformation with the scaling ratio $r = 1/3$. In this case, the resulting curve is the union of 4 non-overlapping curves which is congruent to the curve obtained from the original one by similarity transformation. Therefore, $d_{F,ss}$ of this curve is determined to be $d_{F,S} = \ln 4 / \ln 3 \cong 1.26$ by Eq. (1).

The set S is statistically self-similar when S is the union of N distinct subsets each of which is scaled down by r from the original and is congruent to $r(S)$ in all statistical respects. For

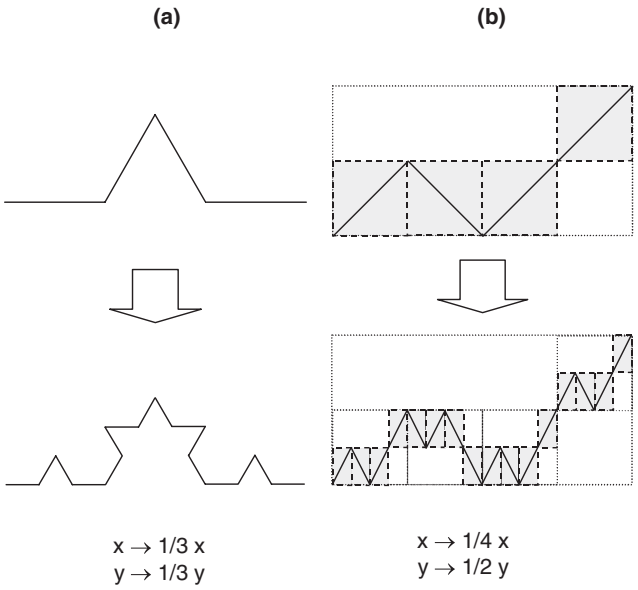


Figure 2. (a) A deterministic self-similar fractal, i.e., the triadic Koch curve, generated by the similarity transformation with the scaling ratio r = 1/3 and (b) a deterministic self-affine fractal generated by the affine transformation with the scaling ratio vector $\mathbf{r}$ = (1/4, 1/2).

such sets the box-counting method[1,61,62,65] is useful in estimating $d_{\mathrm{F,ss}}$ of the set and the resulting box dimension $d_{\mathrm{F,B}}$ coincides with $d_{\mathrm{F,S}}$.

2. Self-Affine Fractal[61,62]

An affine transformation transforms a set of points S at position $\mathbf{x} = (x_1, ..., x_{\mathrm{E}})$ in Euclidean E-dimensional space into a new set of points $\mathbf{r}(S)$ at position $\mathbf{x}' = (r_1 x_1, ..., r_{\mathrm{E}} x_{\mathrm{E}})$ with the different

scaling ratios $0 < r_1,...,r_E < 1$. A bounded set S is self-affine with respect to a scaling ratio vector $\mathbf{r} = (r_1,...,r_E)$ if S is the union of N non-overlapping subsets $S_1,...,S_N$, each of which is congruent to the set $\mathbf{r}(S)$. Figure 2b shows a deterministic self-affine fractal generated by the affine transformation with the scaling ratio vector $\mathbf{r} = (1/4,1/2)$. In this case, the resulting curve is the union of 4 non-overlapping curves which are congruent to the curve obtained from the original one by the affine transformation.

In a similar way, the set S is statistically self-affine when S is the union of N non-overlapping subsets each of which is scaled down by $\mathbf{r}$ from the original and is congruent to $\mathbf{r}(S)$ in all statistical respects.

In contrast to the self-similar case, the self-affine fractal dimension $d_{F,sa}$ of even the simplest self-affine fractal is not uniquely defined.[61,62,66] The similarity dimension $d_{F,S}$ exists only for the self-similar fractal. For the self-affine fractal, there are two different dimensions, i.e., a global dimension and a local dimension.[11,61,62,65-67] The global dimension, which is observed above a certain crossover scale, is simply $d_E - 1$. Here, d_E represents the topological dimension of the Euclidean space where the set is embedded. This means that, at large size level, the self-affine fractal looks essentially smooth.

It is the local dimension that describes the irregularity of the self-affine fractal. The local dimension can be determined by such methods as the box-counting method[1,61,62,65] and the divider-walking method.[61,66] The box dimension $d_{F,B}$ is defined by the Hurst exponent H according to following equation

$$d_{F,B} = d_E - H \qquad 0 < H < 1 \tag{2}$$

and the divider dimension $d_{F,D}$ is given by

$$d_{F,D} = \frac{1}{H} \qquad 0 < H < 1 \tag{3}$$

where H represents the Hurst exponent which is a power exponent observed in the power law between the root mean square (rms) roughness σ_{rms} and the horizontal length L of the self-affine fractal. From Eqs. (2) and (3), it is noted that $d_{\text{F,B}}$ and $d_{\text{F,D}}$ do not coincide in value for the self-affine fractal. Therefore, the Hurst exponent H uniquely characterizes the self-affine fractal.

III. CHARACTERIZATION OF ROUGH SURFACES AND INTERFACES BASED UPON FRACTAL GEOMETRY: METHODS NEEDED FOR THE DETERMINATION OF THE SURFACE FRACTAL DIMENSION

1. Physical Methods

(*i*) *Image Analysis Methods*

The surface roughness is fundamental for several physical properties and physical phenomena on solid surfaces and interfaces. In order to quantify the surface roughness, the solid surfaces and interfaces are investigated using surface profiler or imaging equipments such as scanning electron microscope (SEM), transmission electron microscope (TEM) and scanning probe microscope (SPM). And then the statistical roughness parameters are determined from the digitized surface profile or image. The best known roughness parameter is the root mean square roughness (rms roughness) σ_{rms}, but this parameter is rather inadequate to provide a complete description of three-dimensional surface since it describes only vertical amplitude in a certain lateral direction.

Surface fractal dimension is an adequate roughness parameter since it represents both the vertical and the lateral information of three-dimensional surfaces over a significant range of length scales. In fractal analysis of the three-dimensional digitized image of the surface morphology, SPM is a more adequate technique than SEM and TEM because of its high three-dimensional resolution and nondestructive character. For example, by using a scanning tunneling microscope (STM) and an atomic force

microscope (AFM), the surface morphology is directly described by the digitized height data with a resolution down to the atomic scale, whereas using SEM and TEM, the two-dimensional cross-section of the surface morphology can be described by binary digits.

Several algorithms have been used to determine surface fractal dimension from SPM images. The most popular ones are the power-spectrum method,[1,2,68-71] the triangulation method,[10,37,43,45,61,66,72-75] the perimeter-area method,[18-21,23,25,27,61,62,71,76-78] the structure function method,[2,56,57,70,76,79,80] the variance method,[56,57] and the box-counting method.[1,61,62,65,81,82] Among these algorithms, the triangulation method and the perimeter-area method are introduced in this article as useful and reliable tools to determine the self-similar and the self-affine fractal dimensions, respectively.

(*a*) *Triangulation method*[10,37,43,45,61,66,72-75]

In order to characterize the three-dimensional self-similar fractal surface, the self-similar fractal dimension $d_{\mathrm{F,ss}}$ has been determined from the analysis of the three-dimensional digitized SPM image by using the triangulation method. This method is exactly analogous to the Richardson method[1] for a profile.

Figure 3 describes schematically the algorithm used for determination of $d_{\mathrm{F,ss}}$ by the triangulation method. The square (x,y) plane with a cell size L^2 is first divided into N^2 equal squares. This defines the location of the vertices of a number of triangles. Then the electrode surface is covered by $2N^2$ triangles inclined at various angles with respect to the (x,y) plane. These $2N^2$ triangles have equal projected triangle size TS ($= L / N$), although their real areas are different.

The scaled surface area SSA, i.e., the measured surface area by $2N^2$ triangles, is estimated to be the sum of the areas of all the $2N^2$ triangles. This measurement is iterated as decreasing projected triangle size TS, until every pixel in the AFM image serves as the vertices of the $2N^2$ triangles. Then, $d_{\mathrm{F,ss}}$ of the surface is given by

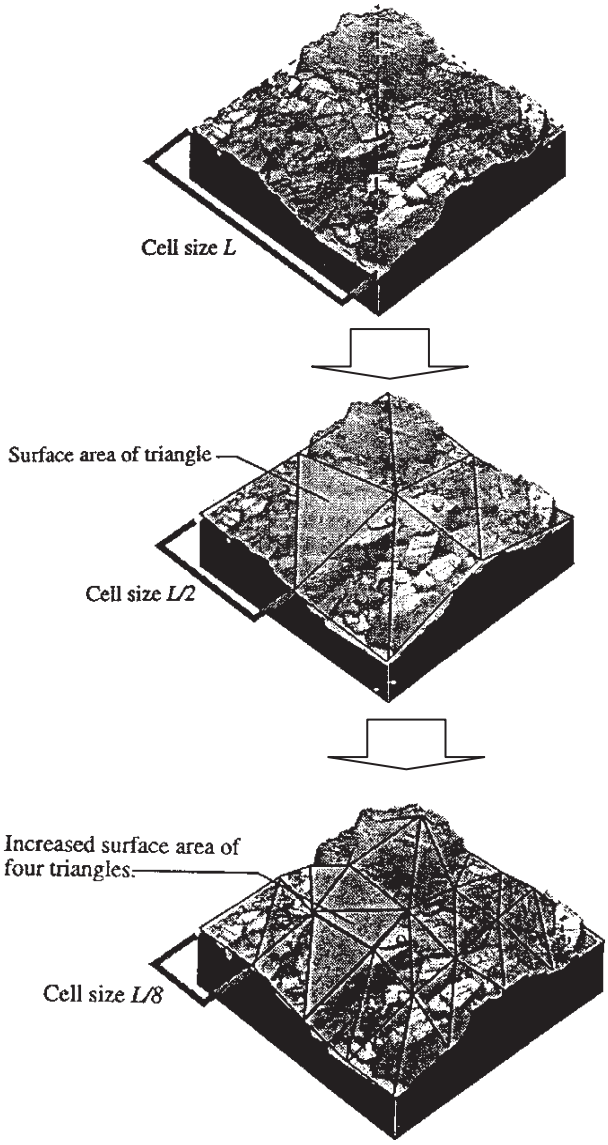


Figure 3. Process of determination of the self-similar fractal dimension of the three-dimensional surface by the triangulation method.

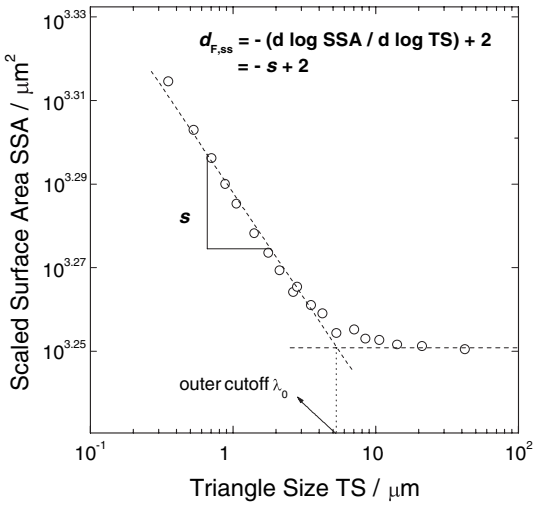


Figure 4. Dependence of the scaled surface area SSA on the projected triangle size TS on a logarithmic scale obtained from the three-dimensional AFM image of the hot-rolled carbon steel surface. Here, s means $\mathrm{d}\ln \mathrm{SSA} / \mathrm{d}\ln \mathrm{TS}$.

$$d_{\mathrm{F,ss}} = -\frac{\mathrm{d}\log \mathrm{SSA}}{\mathrm{d}\log \mathrm{TS}} + 2 \tag{4}$$

After measuring SSAs for the various projected TSs, the resulting SSAs are plotted as a function of the projected TS on a logarithmic scale. Figure 4 gives dependence of the scaled surface area SSA on the projected triangle size TS on a logarithmic scale obtained from the three-dimensional AFM image of the hot-rolled carbon steel surface. From the linear relationship between the logarithm of SSA and the logarithm of projected TS, $d_{\mathrm{F,ss}}$ of the surface is determined using Eq. (4). In addition, it is possible to define a length-scale range where the rough surface shows the

fractal behavior by determining the inner and outer cutoffs in Figure 4. In the case of Figure 4, the spatial outer (upper) cutoff λ_o can be evaluated to be the value of TS where the linear line with the slope of s and the horizontal line intersect. However, the spatial inner (lower) cutoff λ_i can not be evaluated due to the limitation in the resolution of the equipment, e.g., AFM.

(*b*) *Perimeter-area method*[18-21,23,25,27,61,62,71,76-78]

In order to characterize the self-affine fractal surface, the self-affine fractal dimension $d_{F,sa}$ has been determined by using the perimeter-area method. This method is based on the fact that the intersection of a plane with a self-affine fractal surface generates self-similar lakes. This method is available for the images obtained by using SEM and TEM as well as SPM because it is used to analyze the two-dimensional digitized cross-section of the surface. It is well known[61,76] that the area A and the perimeter P of the self-similar lakes are related to their fractal dimension $d_{F,ss}^{L}$ by

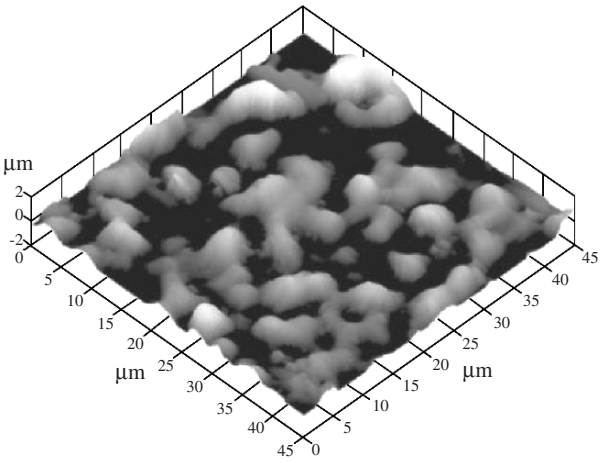
$$P = \beta\, d_{F,ss}^{L}\, A^{d_{F,ss}^{L}/2} \tag{5}$$

where β is a proportionality constant. The self-affine fractal dimension of the original surface $d_{F,sa}$ is related to $d_{F,ss}^{L}$ by

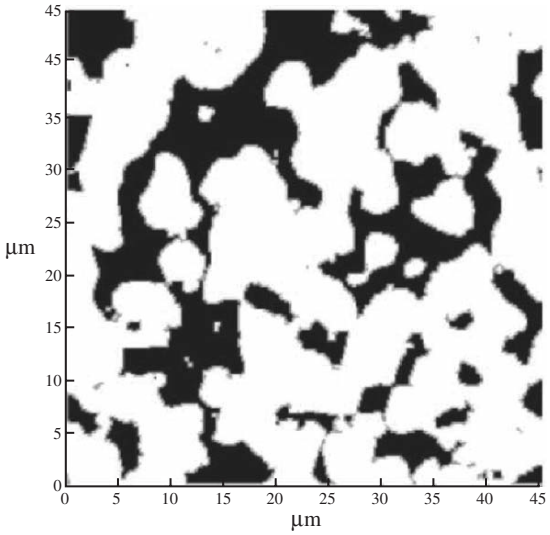
$$d_{F,sa} = d_{F,ss}^{L} + 1 \tag{6}$$

The algorithm used for determination of the self-affine fractal dimension by the perimeter-area method is given as follows:[43,45,77]

The three-dimensional digitized image of the surface morphology is cross-sectioned by a plane and then P and A of the resulting lakes are calculated. For instance, Figure 5a and 5b depict the three-dimensional AFM image of the hot-rolled carbon steel surface filled with water (black pixels) up to a height corresponding to 40 % of the maximum height of the surface and the corresponding two-dimensional description of P (gray pixels)



(a)



(b)

Figure 5. (a) Three-dimensional AFM image of the hot-rolled carbon steel surface filled with water (black pixels) up to a height corresponding to 40% of the maximum height and (b) corresponding two-dimensional description of the perimeters (gray pixels) and areas (gray and black pixels) of the lakes.

along with A (black pixels + gray pixels) of the self-similar lakes, respectively. Here, the gray pixels in Figure 5b are defined as the black pixels neighboring white pixels. P and A of each lake are estimated from the numbers of gray pixels and of both black pixels and gray pixels for each lake, respectively.

After measuring P and A of each self-similar lake, P is plotted as a function of A on a logarithmic scale in Figure 6. The plot of log P vs. log A gave a good linear relation above the threshold area $A_T \approx 2.6\text{x}10^{-13}$ m^2. From the linear relationship between the logarithm of P and the logarithm of A which appears above A_T in Figure 6,

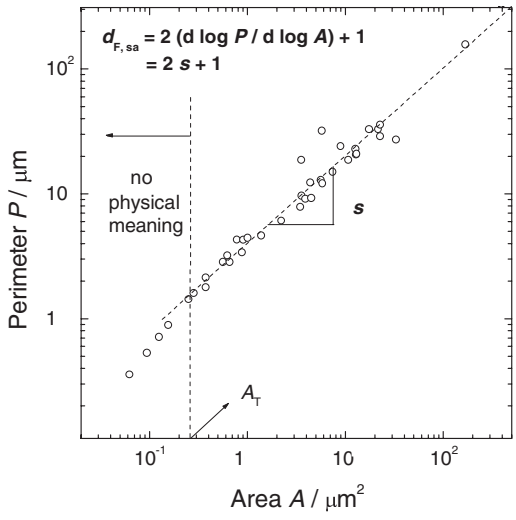


Figure 6. Dependence of the perimeter P on the area A for the self-similar lakes generated by the intersection of the three-dimensional AFM image of the hot-rolled carbon steel with a plane at a height corresponding to 40% of the maximum height. Here, s means d log P/d log A.

the self-similar dimension of the two-dimensional lakes $d_{\mathrm{F,ss}}^{\mathrm{L}}$ is determined using Eq. (5) and then the self-affine fractal dimension of the three-dimensional surface $d_{\mathrm{F,sa}}$ is determined using Eq. (6) as well. It is noted that another linear relation below A_{T} is physically meaningless due to the limitation of the AFM measurement.[20]

(ii) Molecular Probe Method Using Gas Adsorption

Since Avnir and Pfeifer's pioneer works[83-86] regarding the characterization of the surface irregularity at the molecular level by applying the fractal theory of surface science, the molecular probe method using gas adsorption has played an important role in the determination of surface fractal dimension of the porous and particulate materials.

There are two molecular probe methods available for the determination of surface fractal dimension. One is the multiprobe method (MP method),[83,84,87–100] which uses several kinds of multiprobe molecules with different molecular sizes and requires the number of adsorbed molecules to form a monolayer N_{mono} for each probe molecule. If the probe molecule is varied through a series of spheres with radius r_{m}, the surface fractal dimension is given by Eq. (7):

$$N_{\mathrm{mono}} \propto r_{\mathrm{m}}^{-d_{\mathrm{F,MP}}} \qquad 2 \le d_{\mathrm{F,MP}} \le 3 \tag{7}$$

Here, the surface fractal dimension determined by using the MP method is denoted by $d_{\mathrm{F,MP}}$. The probe molecules need not be spherical, provided they belong to a homologous series for which the ratio $[\mathrm{linear\,extent}\,r]^2$ to effective cross-section area A_{c} is the same for all members, i.e., an isotropic series. In that case, Eq. (7) turns into

$$N_{\mathrm{mono}} \propto A_{\mathrm{c}}^{-\frac{d_{\mathrm{F,MP}}}{2}} \qquad 2 \le d_{\mathrm{F,MP}} \le 3 \tag{8}$$

The MP method is widely applicable to various solid surfaces in whose system N_{mono} can be accurately determined. However, the MP method can not be routinely applied to the microporous system due to a serious problem in the concept of monolayer adsorption for the microporous system.[101] This problem is whether the monolayer formation on the surface of the micropore is eligible or not.

Recently, in the theoretical studies on the simulation for N_2 adsorption in micropore, some researchers[102–104] reported that the monolayer adsorption occurs even in the micropore whose pore width is greater than the bilayer thickness of N_2 (about 0.7 nm). In addition, Kaneko et al. showed the presence of the orientational phase transition of N_2 on the graphitic micropore wall, which is the same as the phase transition of the monolayer on the flat graphite surface,[105] and gave an effective method for the surface area determination in the microporous system.[106] Therefore, even for micropores whose width is greater than 0.7 nm, $d_{F,MP}$ can be determined using Eq. (8), if the monolayer evaluation fit for the micropore is introduced. Based upon these theoretical and experimental results, Sato *et al*. determined the values of $d_{F,MP}$ of less-crystalline carbon micropore walls using Eq. (8).[107]

The other molecular probe method is the single-probe method (SP method), which is separately proposed by Avnir and Jaroniec,[93] and Pfeifer *et al*.[108–112] In the SP method, a single adsorption isotherm is analyzed using a modified FHH theory. The FHH model was developed independently by Frenkel,[113] Halsey,[114] and Hill,[115] and describes the multilayer adsorption coverage. Since the SP method uses only one probe molecule, this method is more convenient than the MP method. However, there are many theoretical limitations in applying the SP method to determination of the surface fractal dimension. Therefore, it is really necessary to discuss whether the SP method is an adequate tool to investigate the surface fractal dimension or not before applying the SP method to certain system.

The theoretical assumptions used in the SP method are as follows: (1) The detailed structure of the layer of the adsorbed molecules is not important and is regarded as a liquid film. (2)

The thickness of the layer of the adsorbed molecules is the characteristic distance scale for fractal surface. (3) Van der Waals attraction forces between solid/gas interactions and the liquid/gas surface tension forces are contributed to the grand potential of the system.

On the basis of the theory of Pfeifer *et al.*,[108–112] the surface fractal dimension can be determined from the gas adsorption isotherms according to the following equation:

$$\ln\left(\frac{V}{V_{\text{mono}}}\right) = const. + C\left[\ln\ln\left(\frac{p^0}{p}\right)\right] \tag{9}$$

where V is the volume of adsorbed gas molecules; V_{mono}, the volume of monolayer coverage; C, the power law exponent which is dependent on the surface fractal dimension determined by using the SP method $d_{\text{F,SP}}$; p^0, the saturation pressure of the gas and p represents the adsorption equilibrium pressure of the gas.

At the early stage of multilayer build-up, the interface is dominantly controlled by van der Waals forces between solid/gas interactions and the liquid/gas surface tension forces are negligible. Then, the relationship between C and $d_{\text{F,SP}}$ is written as:

$$C = \left(\frac{d_{\text{F,SP}} - 3}{3}\right) \tag{10}$$

For higher coverage, the interface is controlled by the liquid-gas surface tension forces (capillary condensation) and van der Waals forces between solid/gas interactions are negligible. Then, the relationship between C and $d_{\text{F,SP}}$ changes to the following expression:

$$C = \left(d_{\text{F,SP}} - 3\right) \tag{11}$$

Van der Waals forces between solid/gas interactions and the liquid/gas surface tension forces represent the limiting cases, but in general both the forces competitively affect the adsorption process. Therefore, in determining the surface fractal dimension of the carbon specimen, it is very important to use appropriate relation between C and $d_{F,SP}$. According to Ismail and Pfeifer,[111] the threshold for the dominant forces between van der Waals forces and the liquid/gas surface tension forces is given as:

$$\varepsilon = 3\,(1 + C) - 2 \tag{12}$$

If $\varepsilon < 0$, the liquid/gas surface tension forces are dominant, while the van der Waals forces are dominant if $\varepsilon \geq 0$.

It is reported[116,117] that as more adsorbed layers are built up, the interface between the adsorbent and the adsorbed molecules becomes smooth, and hence the surface fractal dimension would no longer describe the interface but would describe the adsorbed molecule agglomerates. Also, Eq. (9) is only valid when the adsorbed layer exceeds monolayer coverage. Therefore, for the correct calculation, $d_{F,SP}$ should be determined from the linear segment in the range between the monolayer coverage and the multiple layer coverage just before the smoothing effect appears.

In addition, the number of adsorbed molecule layers n can be calculated from the following relationship.[111]

$$n = \frac{t_{ad}}{a_0} = \left(\frac{V}{V_{mono}}\right)^{1/(3-d_{F,SP})} \tag{13}$$

Here, the actual thickness of adsorbed molecule layers t_{ad} is obtained by multiplying n by adsorbed molecule diameter a_0. The actual thickness range in which the fractal geometry is satisfied represents the length-scale cutoff range of fractality.[116,118] Here, the value of the lower (inner) cutoff length $t_{ad,min}$ is equal to that value of monolayer thickness and that of the upper (outer) cutoff length $t_{ad,max}$ can be evaluated by using following equation which is obtained from Eq. (13):

$$t_{\mathrm{ad,max}} = t_{\mathrm{ad,min}} \left(\frac{V_{\max}}{V_{\min}} \right)^{1/\left(3-d_{\mathrm{F,SP}}\right)} \tag{14}$$

where $V_{\min}$ and $V_{\max}$ are the lower and upper limits of the power-law regimes in the plot of $\ln V$ against $\ln \ln(p^0 / p)$.

2. Electrochemical Methods

(*i*) *Potentiostatic Current Transient Method*

Consider an inert metal with a fractal electrode immersed into an electrolyte containing a redox couple. We assume semi-infinite diffusion of the oxidized species Ox and the reduced species Red in the electrolyte. For the sake of simplicity we assume that the solution contains initially (at the time, $t = 0$) only the oxidized form and the bulk and surface concentrations are identical, i.e., $c_{\mathrm{ox}}^{\mathrm{b}} = c_{\mathrm{ox}}^{\mathrm{s}}$. The electrode is initially subjected to an electrode potential where no reaction takes place. The only reaction occurring when the potential is lowered, is the reversible reduction of Ox to Red, i.e., $Ox + ze = Red$. In addition, it is assumed that the overall reaction is limited by diffusion of Ox in electrolyte.

When a potential step is applied at a flat electrode, the time dependence of Faradaic current I is well described by the Cottrell equation[119]

$$I(t) = \frac{zFA_{\mathrm{ex}} \widetilde{D}_{\mathrm{Ox}}{}^{1/2} c_{\mathrm{Ox}}^{\mathrm{b}}}{\pi^{1/2}} t^{-1/2} \tag{15}$$

where z is the number of electrons transferred per the oxidized species in the redox reaction; F, the Faraday constant; A_{ex}, the exposed surface area of the working electrode and $\widetilde{D}_{\mathrm{Ox}}$ represents the chemical diffusivity of the oxidized species.

In the case of a fractal electrode, the relationship between the current and the time follows generalized Cottrell equation derived

by Nyikos and Pajkossy based upon the scaling law for current[120,121]

$$I(t) = \sigma_g t^{-\alpha} = \sigma_g t^{-(d_F - 1)/2} \tag{16}$$

where d_F is the fractal dimension of the electrode; α, the power exponent and σ_g is the generalized Cottrell coefficient defined by following equation.[121]

$$\begin{aligned} \sigma_g &= \frac{zFA_{micr}c_{Ox}^b \tilde{D}_{Ox}^{1/2}}{\pi^{1/2}} \left[\frac{\gamma\lambda_o^2}{\tilde{D}_{Ox}} \right]^{(d_F/2)-1} \\ &= \frac{zFA_{macr}c_{Ox}^b \tilde{D}_{Ox}^{1/2}}{\pi^{1/2}} \left[\frac{\gamma\lambda_i^2}{\tilde{D}_{Ox}} \right]^{(d_F/2)-1} \end{aligned} \tag{17}$$

Here, A_{micr} is the microscopic area which is defined within the fractal limits; γ, the dimensionless geometrical parameter; λ_o, the spatial outer (upper) cutoff; A_{macr}, the macroscopic area which is defined beyond the fractal limits ($= A_{ex}$ in Eq. 15) and λ_i represents the spatial inner (lower) cutoff.

The power-law behavior represented by Eq. (16) is valid between the spatial cutoffs λ_i and λ_o corresponding to the temporal inner (shorter) cutoff τ_i and the temporal outer (longer) cutoff τ_o. So far as the diffusion process is concerned, the diffusion layer length acts as a yardstick length.[122] Therefore, the relationship between the temporal and spatial cutoffs is given as[121]

$$\lambda = \left(\frac{\tau\tilde{D}_{Ox}}{\gamma} \right)^{1/2} \tag{18}$$

When d_F is around 2.5, γ is about $1/\pi^2$, and when d_F is near 2.0, γ is assumed to be about $1/\pi$.[121]

With the help of fractional calculus, Dassas and Duby[123] have worked on the problem of diffusion towards the fractal interfaces. They have proposed the following generalized diffusion equation involving a fractional derivative operator:

$$\frac{\partial^{(3-d_F)}c}{\partial t^{(3-d_F)}} = \widetilde{D}^{*}_{Ox}\frac{\partial^2 c}{\partial x^2} \tag{19}$$

Here, $\widetilde{D}^{*}_{Ox}$ is the fractional diffusivity defined as $K^{(4-2d_F)}R_o^{(2d_F-4)}\widetilde{D}_{Ox}^{(3-d_F)}$ (K is a constant related to the fractal dimension and R_o is the side length of a square electrode), and $\partial^{\nu}/\partial t^{\nu}$ means the Riemann-Liouville mathematical operator of fractional derivative:

$$\frac{\partial^{\nu}c}{\partial t^{\nu}} = \frac{1}{\Gamma(1-\nu)}\frac{d}{dt}\int_0^t \frac{c(\xi)}{(t-\xi)^{\nu}}d\xi \tag{20}$$

During the potential step experiment, the initial condition (I.C.) and the boundary condition (B.C.) are given as

I.C.: $c = c_{Ox}^{b}$ for $0 \le x < \infty$ at $t = 0$ (21)

B.C.: $c = c_{Ox}^{b}$ for $x \to \infty$ at $t \ge 0$ (semi-infinite constraint) (22)

$c = 0$ for $x = 0$ at $t > 0$ (potentiostatic constraint) (23)

From the Laplace transforms of Eqs. (19) and (21) to (23), a generalized Cottrell equation which describes the current response on the applied potential step is obtained as following equation

$$I(t) = \frac{zFA_{ex}\widetilde{D}^{*\,1/2}_{Ox}c_{Ox}^{b}}{\Gamma\left(\frac{3-d_F}{2}\right)}t^{-\alpha} = \frac{zFA_{ex}\widetilde{D}^{*\,1/2}_{Ox}c_{Ox}^{b}}{\Gamma\left(\frac{3-d_F}{2}\right)}t^{-\left(\frac{d_F-1}{2}\right)} \tag{24}$$

where $\Gamma\left(\frac{3-d_{\mathrm{F}}}{2}\right)$ is the gamma function of $\left(\frac{3-d_{\mathrm{F}}}{2}\right)$. The Cottrell equation for flat diffusion is retrieved for $d_{\mathrm{F}} = 2$.

(ii) Linear Sweep/Cyclic Voltammetry

We consider again the redox reaction $Ox + ze = Red$ with a solution initially containing only the oxidized form Ox. The electrode is initially subjected to an electrode potential E_{i} where no reaction takes place. For the sake of simplicity, it is assumed that the diffusion coefficients of species Ox and Red are equal, *i.e.,* $\widetilde{D} = \widetilde{D}_{\mathrm{Ox}} = \widetilde{D}_{\mathrm{Red}}$. Now, the potential E is linearly increased or decreased with $E(t) = E_{\mathrm{i}} \pm \nu t$ (ν is a potential scan rate, and signs "+" and "–" represent anodic scan and cathodic scan, respectively.) Under the assumption that the redox couple is reversible, the surface concentrations of Ox and Red, i.e., $c^{\mathrm{s}}_{\mathrm{Ox}}$ and $c^{\mathrm{s}}_{\mathrm{Red}}$, respectively, are always determined by the electrode potential through the Nernst equation

$$E = E_{1/2} + \frac{RT}{zF}\ln\left(\frac{c^{\mathrm{s}}_{\mathrm{Ox}}}{c^{\mathrm{s}}_{\mathrm{Red}}}\right) \tag{25}$$

where $E_{1/2}$ means the half-wave potential, i.e., the potential bisecting the distance between anodic and cathodic peaks in a cyclic voltammogram. In addition, it is assumed that the overall reaction is limited by diffusion of Ox or Red in electrolyte.

When the potential is varied at the flat electrode, the peak current I_{peak} is related to ν as the Randles-Sevcik equation[119]

$$I_{\mathrm{peak}} = zFA_{\mathrm{ex}}c^{\mathrm{b}}_{\mathrm{Ox}}(\pi\widetilde{D}_{\mathrm{Ox}})^{1/2}\left(\frac{zF}{RT}\nu\right)^{1/2}\chi_{\max}\left(\frac{zF}{RT}\nu t\right) \tag{26}$$

where R is the gas constant; T, the temperature and $\chi_{\max}$ represents the maximum value of the dimensionless parameter χ.

It is interesting to note that the conventional Randles-Sevcik equation, relating I_{peak} to $\nu^{1/2}$, does not hold for the fractal electrode. Instead it is replaced with the following generalized Randles-Sevcik equation, relating I_{peak} to $\nu^{(d_{\text{F}}-1)/2}$.[121,124-126]

$$I_{\text{peak}} = \frac{zFA_{\text{macr}}\widetilde{D}_{\text{Ox}}^{1/2}\left(\dfrac{\gamma\lambda_{\text{i}}^2}{\widetilde{D}_{\text{Ox}}}\right)^{(d_{\text{F}}/2)-1} c_{\text{Ox}}^{\text{b}}\Gamma\left(\dfrac{3-d_{\text{F}}}{2}\right)\Gamma\left(\dfrac{d_{\text{F}}-1}{2}\right)}{\Gamma\left(\dfrac{1}{2}\right)} \times\left(\frac{zF}{RT}\nu\right)^{(d_{\text{F}}-1)/2}\chi_{\max}\left(\frac{zF}{RT}\nu t\right) \tag{27}$$

According to Eq. (27), Strømme et al.[125,126] developed systematically the peak-current method to determine the fractal dimension of the electrode surface by using cyclic voltammetry. It must be recalled that this method is valid when the recorded current is limited by diffusion of the electroactive species to and away from the electrode surface. Since the distribution of the reaction sites provides extensive information about the surface geometry, the fractal dimension of the reaction site distribution may agree with the fractal dimension of the electrode surface which is completely electrochemical-active. In addition, it is well known that this method is insensitive to the IR drop in the electrolyte.[126]

As mentioned in potentiostatic current transient method, when the fractal dimension is determined by using diffusion-limited electrochemical technique, the diffusion layer length acts as a yardstick length.[122] In the case of cyclic voltammetry, it was derived as follows[127,128]

$$\Delta\lambda \approx \frac{zFA_{\text{macr}}\widetilde{D}_{\text{Ox}}\, c_{\text{Ox}}^{\text{b}}}{I_{\text{peak}}} \tag{28}$$

where $\Delta\lambda$ is the diffusion layer length.

Therefore, we can define a length scale range where the rough surface shows the fractal behavior by determining the inner and outer cutoffs using Eq. (28).[121] In other words the fractal dimensions determined from the power relation between I_{peak} and v in the cyclic voltammetry have a physical meaning in the range between the inner cutoff and the outer cutoff.

If the capacitive current exists significantly in redox system, it is necessary to remove the effect of capacitive current on I_{peak} for the accurate determination of d_F. In addition, since this effect is larger with increasing potential scan rate, regardless of the surface roughness of the electrode surface, the inner cutoff determined by using Eq. (28) might include some ambiguity.

(iii) Ac Impedance Spectroscopy

A constant phase element (CPE) rather than the ideal capacitance is normally observed in the impedance of electrodes. In the absence of Faradaic reactions, the impedance spectrum deviates from the purely capacitive behavior of the blocking electrode, whereas in the presence of Faradaic reactions, the shape of the impedance spectrum is a depressed arc. The CPE shows power law frequency dependence as follows[129,130]

$$Z(\omega) = \frac{1}{B}(j\omega)^{-\alpha_{CPE}} \tag{29}$$

where B means the CPE coefficient; j, the imaginary unit; α_{CPE}, the CPE exponent; Z, the impedance and ω represents the angular frequency.

Considerable effort has been made to understand the origin of the CPE. One of the most possible causes of the CPE is of geometric origin: an irregular and porous electrode geometry causes current density inhomogeneities and thus yields deviations from ideal behavior. A number of theoretical papers[131-144] have investigated the relationship between the fractal geometry of the

electrode and the CPE exponent α_{CPE}. Pajkossy and Nyikos[144] have shown that for a self-similar fractal working electrode with $d_{\mathrm{F}} > 2.0$, and when the counter electrode is located far from the working electrode, the relationship between d_{F} and α_{CPE} is given as

$$\alpha_{\mathrm{CPE}} = \frac{1}{d_{\mathrm{F}} - 1} \tag{30}$$

However, the relationship between d_{F} and α_{CPE} is dependent on the model used in the fractal characterization of the electrode.[144,145]

Furthermore, it has been reported that the CPE originated from the surface inhomogeneities and adsorption effects rather than the irregular geometry because the CPE due to irregular geometry appears at much higher frequencies than is usual in electrochemical methodologies.[145-147] Therefore, for the reliable fractal characterization of electrode surface in terms of α_{CPE}, it must be rigorously checked whether the surface state of the electrode is homogeneous or not and whether the effect of the surface roughness on the CPE dominates the surface inhomogeneities and adsorption effects or not.

The length scales corresponding to the frequency ranges where the fractal scaling property is expected in the impedance measurements can be considered. According to Pajkossy,[145] the relationship between the length scale λ and the frequency f is give as

$$f \approx \frac{1}{2\pi C_{\mathrm{e}} \rho_{\mathrm{el}} \lambda} \tag{31}$$

where C_{e} means the electrode capacitance per unit area and ρ_{el} is the specific resistivity of the electrolyte.

IV. INVESTIGATION OF DIFFUSION TOWARDS SELF-AFFINE FRACTAL INTERFACE

Since diffusing species move randomly in all directions, the diffusing species may sense the self-affine fractal surface and the self-similar fractal surface in quite different ways. Nevertheless a little attention has been paid to diffusion towards self-affine fractal electrodes. Only a few researchers have realized this problem: Borosy et al.[148] reported that diffusion towards self-affine fractal surface leads to the conventional Cottrell relation rather than the generalized Cottrell relation, and Kant[149,150] discussed the anomalous current transient behavior of the self-affine fractal surface in terms of power spectral density of the surface.

Recently, Pyun et al.[43,45] gave a clear interpretation of diffusion towards self-affine fractal interface. They investigated theoretically how the diffusing ions sense the self-affine fractal interface during the diffusion-limited process[43] and then provided successful experimental evidence of the theoretical investigation.[45] Here, let us explore their works in detail.

In their theoretical work,[43] the various self-affine fractal interfaces were mathematically constructed employing the Weierstrass function $f_{\mathrm{WS}}(x)$.[151,152]

$$f_{\mathrm{WS}}(x) = \sum_{k=1}^{N} a^k \cos(b^k x) \tag{32}$$

where $0 < a < 1 < b$ with $ab \geq 1$. The curve $f_{\mathrm{WS}}(x)$ is known to have the following value for the self-affine fractal dimension $d_{\mathrm{F,sa}}$:

$$d_{\mathrm{F,sa}} = 2 + \frac{\ln a}{\ln b} \tag{33}$$

A self-affine fractal profile with $d_{\mathrm{F,sa}} = 1.5$ was obtained by taking a, b, and N as 0.8165, 1.5, and 50, respectively. Next, various affine functions $h(x)$ were made by multiplying $f_{\mathrm{WS}}(x)$

by morphological amplitude η (=0.1, 0.3, 0.5, 1.0, 2.0, and 4.0), *i.e.*, $h(x) = \eta f_{\mathrm{WS}}(x)$. The resulting six self-affine functions are presented in Figure 7. Here, it should be noted that the original self-affine function $f_{\mathrm{WS}}(x)$ and the magnified functions $h(x)$ have the same self-affine fractal dimension $d_{\mathrm{F,sa}} = 1.5$, but they are quite different from each other in the amplitude.

In order to examine the current response to the imposition of the potential step on the self-affine fractal interface, the current transients were calculated theoretically by random walk simulation.[153] The simulation cell was taken as the square area bottom boundary which is replaced by one of the self-affine fractal profiles in Figure 7. The details of the simulation condition were described in their publication.[43]

Figures 8a and 8b present the simulated current transients obtained from the self-affine fractal interfaces of $\eta = 0.1; 0.3; 0.5$ and $\eta = 1.0; 2.0; 4.0$, respectively, embedded by the Euclidean two-dimensional space. It is well known that the current-time relation during the current transient experiment is expressed as the generalized Cottrell equation of Eqs. (16) and (24).[154] So, the power exponent $-\alpha$ should have the value of - 0.75 for all the above self-affine fractal interfaces.

However, the simulated current transients in Figure 8 never exhibited the expected power exponent of –0.75 with the exception of the original self-affine fractal interface ($\eta = 1$): the simulated current transients showed more negative power dependence with increasing η of the interface before the temporal outer (longer) cutoff of fractality,[121,155] followed by the power dependence of - 0.5. In addition, in the case of relatively large amplitudes ($\eta = 2.0$ and 4.0), the current transients were roughly characterized by two-stage power dependence before the temporal outer (longer) cutoff of fractality.

This result strongly indicates that the self-affine fractal dimension $d_{\mathrm{F,sa}}$ is not always sufficient to describe the atomic/ionic diffusion of the self-affine fractal interface and at the same time the morphological amplitude of the interface additionally plays an important role in current transient behavior.

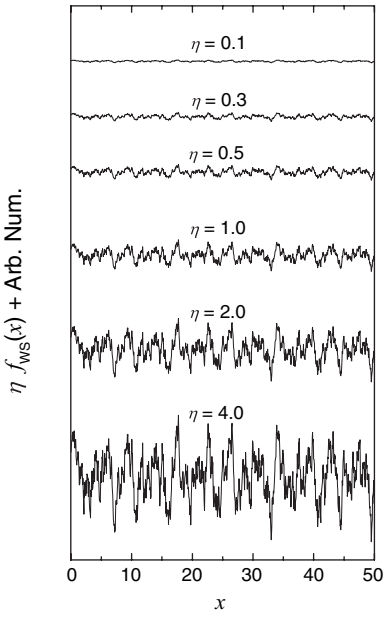


Figure 7. Self-affine fractal profiles of various amplitudes with a self-affine fractal dimension of 1.5, obtained from the Weierstrass function $f_{\mathrm{WS}}(x) = \sum_{k=1}^{N} a^k \cos(b^k x)$, where $0 < a < 1 < b$ with $ab \geq$ a. Reprinted from H. -C. Shin et al., A study on the simulated diffusion-limited current transient of a self-affine fractal electrode based upon the scaling property, *J. Electroanal. Chem.*, **531** p. 101, Copyright © 2002, with permission from Elsevier Science.

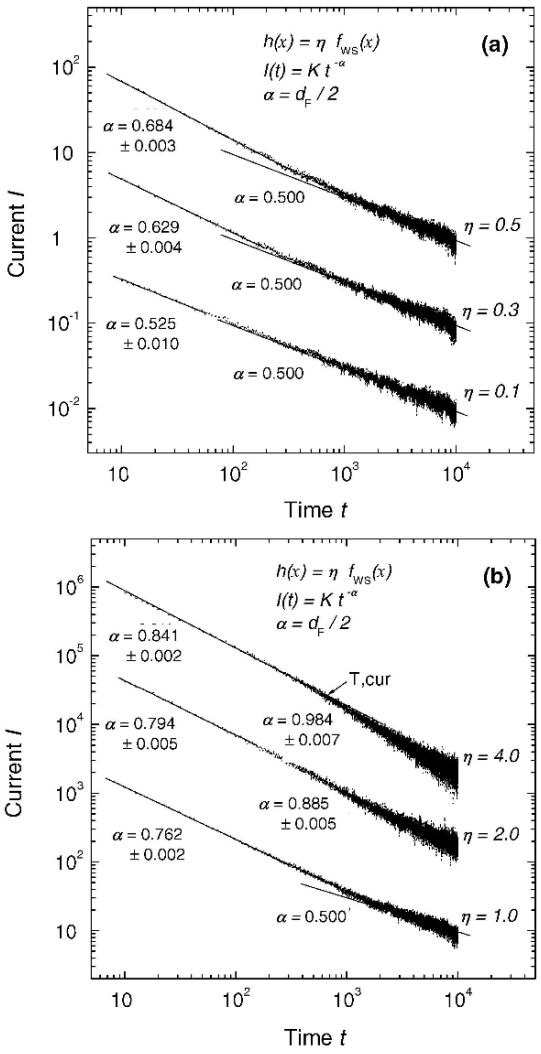


Figure 8. Simulated current transients obtained from the self-affine fractal profiles $h(x)$ of various morphological amplitudes η of: (a) 0.1, 0.3, and 0.5; (b) 1.0, 2.0, and 4.0 in $h(x) = \eta f_{WS}(x)$. Reprinted from H.-C. Shin et al., A study on the simulated diffusion-limited current transient of a self-affine fractal electrode based upon the scaling property, *J. Electroanal. Chem.*, **531,** p. 101, Copyright © 2002, with permission from Elsevier Science.

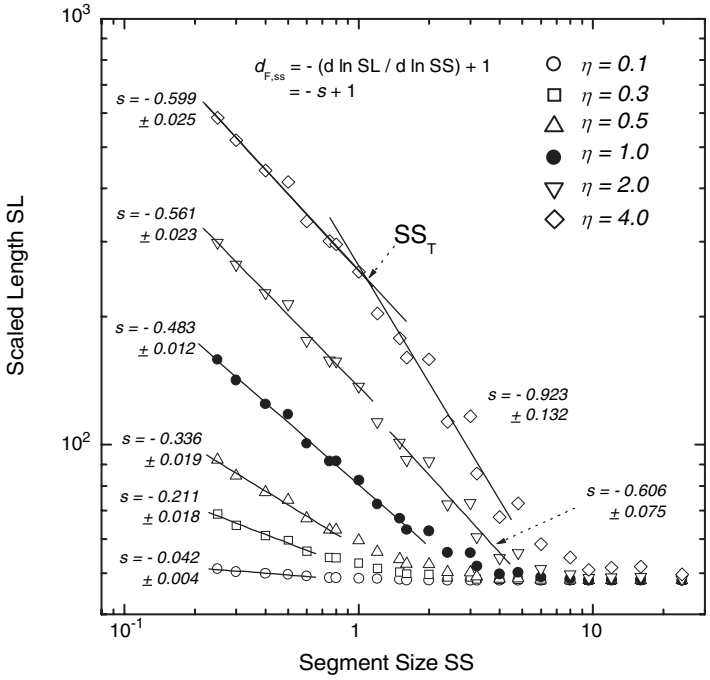


Figure 9. Dependence of the scaled length SL on the projected segment size SS on a logarithmic scale obtained from the self-affine fractal profiles in Figure 7 by using the triangulation method for Euclidean two-dimensional space. Here, s means d ln SL / d ln SS. Reprinted from H. -C. Shin et al., A study on the simulated diffusion-limited current transient of a self-affine fractal electrode based upon the scaling property, *J. Electroanal. Chem.* **531**, p. 101, Copyright © 2002, with permission from Elsevier Science.

It should be stressed here that the specific power dependences from the above self-affine fractal interfaces are maintained even during the relatively long time (or number of random jumps) interval. This implies that the morphology of the self-affine fractal interfaces tested is possibly characterized by the self-similar fractal dimension within a relatively wide spatial cutoff range.

Two-stage power dependence at $\eta > 1.0$ further implies that local small irregularity regions, i.e., the regions with high

curvature, have a different self-similar scaling property from local large irregularity regions, i.e., the regions with low curvature, after magnification process: local small irregularity regions seem to be less sensitive to vertical magnification than local large irregularity regions. This can be easily justified in a geometrical point of view by scaling property analysis for magnified curves (see Figure 9). At this point, one could predict two-stage power dependence even at $\eta \leq 1.0$. In this case, the second power exponent should be the value between the first power exponent and - 0.5 according to the above. Unfortunately, the ambiguity of the current transient due to transition of fractality to non-fractality in the vicinity of the possible second power dependence region hinders us and one has to estimate the second power exponent in this region.

Under the assumption that the morphology of the self-affine interface has the self-similar scaling property, the apparent self-similar fractal dimension $d_{F,ss}$ of the electrode was calculated from the current transients in Figure 8 by Eqs. (16) and (24). The resulting fractal dimension is listed in Table 1 (2nd column).

Table 1
Fractal Dimensions of the Profiles $h(x)$ at Various Morphological Amplitudes η in $h(x) = \eta f_{WS}(x)$ Determined by the Current Transient Technique (2nd Column) and the Triangulation Method (3rd Column). Here, $f_{WS}(x)$ Means the Weierstrass Function with a Self-Affine Fractal Dimension $d_{F,sa} = 1.5$[a]

η \ d_F	d_F Determined by the current transient technique	$d_{F,ss}$ Determined by the triangulation method
0.1	1.050 ± 0.020	1.042 ± 0.004
0.3	1.258 ± 0.008	1.211 ± 0.018
0.5	1.368 ± 0.006	1.336 ± 0.019
1.0	1.524 ± 0.004	1.483 ± 0.012
2.0	1.588 ± 0.010; 1.770 ± 0.010	1.561 ± 0.023; 1.606 ± 0.075
4.0	1.682 ± 0.004; 1.968 ± 0.014	1.599 ± 0.025; 1.923 ± 0.132

[a] Reprinted from H. -C. Shin et al., A Study on the Simulated Diffusion-Limited Current Transient of a Self-Affine Fractal Electrode Based Upon the Scaling Property, *J. Electroanal. Chem.* 531, p. 101, Copyright © 2002, with Permission from Elsevier Science.

Now, it is necessary to check the self-similar scaling properties of the original and the magnified self-affine fractal profiles $h(x)$ of Figure 7. For this purpose, the profiles were analyzed by the triangulation method. The analysis of a fractal curve embedded in the Euclidean two-dimensional space by the triangulation method is performed in the following manner: the curve with a width L is first divided into N equal segments (then the segment size SS becomes L/N). Next, the scaled length SL of the curve is calculated by

$$\mathrm{SL} = \sum_{i=0}^{N-1} \left\{ \left[h\left(\frac{L}{N}(i+1)\right) - h\left(\frac{L}{N} i\right) \right]^2 + \left(\frac{L}{N}\right)^2 \right\}^{1/2} \tag{34}$$

The above operation is iterated at various segment sizes. Finally, the self-similar fractal dimension of the profile embedded by the two-dimensional space is given by

$$d_{\mathrm{F,ss}} = -\frac{\mathrm{d}\ln \mathrm{SL}}{\mathrm{d}\ln \mathrm{SS}} + 1 \tag{35}$$

Figure 9 demonstrates the dependence of the scaled length SL on the segment size SS obtained from the self-affine fractal profiles in Figure 7 by using the triangulation method for the Euclidean two-dimensional space. The linear relation was clearly observed for all the self-affine fractal curves, which is indicative of the self-similar scaling property of the curves.

The self-similar fractal dimension determined from Figure 9 by Eq. (35) is listed in Table 1 (3rd column). The values of $d_{\mathrm{F,ss}}$ obtained from the simulated current transients of Figure 8 shared very well with those values determined by the triangulation method. Moreover, the temporal transition point, e.g., $\tau_{\mathrm{T,tri}} \approx 900$ at $\eta = 4.0$, estimated from the spatial transition point, $\mathrm{SS}_{\mathrm{T}} \approx 1.05$ at $\eta = 4.0$, from the first linearity (self-similar region) to the second linearity (self-similar region) in Figure 9 by the approximate relation $\tau_{\mathrm{T}} \approx SS_{\mathrm{T}}^2 / 2D$ (where D is random diffusion coefficient) quantitatively coincided well with that temporal

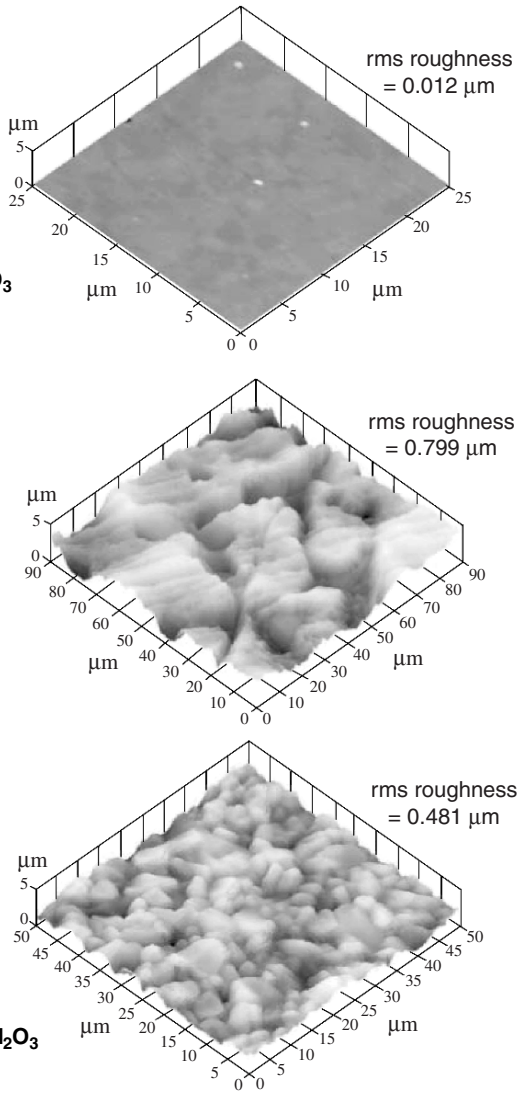
transition point, $\tau_{T,cur} \approx 700$ at $\eta = 4.0$, in the current transients in Figure 8, which justifies their analysis.

From the above results, it is noted that the self-similar scaling property investigated by the triangulation method can be effectively utilized to analyze the diffusion towards the self-affine fractal interface. This is the first attempt to relate the power dependence of the current transient obtained from the self-affine fractal curve to the self-similar scaling properties of the curve.

This theoretical result was experimentally confirmed with the help of AFM and cyclic voltammetry by Go et al.[45] In that work, they first prepared three kinds of self-affine fractal Pt film electrodes by *dc* sputtering of Pt on three kinds of substrates with different roughnesses. And then, surface morphologies of the three electrode specimens were examined by using AFM, and the cyclic voltammograms were obtained from the electrode specimens at various scan rates. Finally, the fractal dimensions of the electrode surfaces were determined from the analyses of the AFM image by both the perimeter-area method and the triangulation method and also from the analysis of the cyclic voltammograms by the peak-current method, and the resulting fractal dimensions were compared with one another.

Figures 10a, 10b, and 10c illustrate the surface morphologies of Pt/polished Al_2O_3, Pt/etched Ni and Pt/unpolished Al_2O_3 electrodes, respectively, investigated with AFM. The three-dimensional AFM image of Pt/polished Al_2O_3 electrode shows a very flat surface. However, the three-dimensional AFM images of Pt/etched Ni electrode and Pt/unpolished Al_2O_3 electrode show the coarse and scratched protrusions with a lateral extent of 20–40 μm and the small and rounded protrusions with a lateral extent of 1–5 μm, respectively. These morphologies of the Pt films are characteristic of the bare surface of the substrates: polished Al_2O_3, etched Ni, and unpolished Al_2O_3.

To quantitatively characterize the surface roughness of the Pt films, the rms roughness σ_{rms} of the electrodes was evaluated and the resulting values are indicated in Figures 10a, 10b and 10c, respectively. From the comparison between σ_{rms} and the scan size in value, it is indicated that the height fluctuation of all the film



(a) Pt/polished Al_2O_3

(b) Pt/etched Ni

(c) Pt/unpolished Al_2O_3

Figure 10. Three-dimensional AFM images of (a) Pt/polished Al_2O_3, (b) Pt/etched Ni, and (c) Pt/unpolished Al_2O_3 electrodes. Reprinted from J. -Y. Go et al., A study on ionic diffusion towards self-affine fractal electrode by cyclic voltammetry and atomic force microscopy, *J. Electroanal. Chem.*, **549**, p. 49,

surfaces (less than one micrometer) is much smaller than the scan size of the film surfaces (several tens micrometer). This result means that the *dc* sputter-deposited Pt film electrodes have self-affine fractal surface rather than self-similar fractal surface.[10]

In order to characterize the self-affine scaling properties of the fractal Pt films, the self-affine fractal dimensions of the film surfaces $d_{F,sa}$ were determined by using the perimeter-area method. Figures 11a, 11b and 11c demonstrate on a logarithmic scale the dependence of the perimeter P on the area A of the lakes at the height corresponding to 40% of the maximum height of the surfaces of Pt/polished Al_2O_3 (Figure 10a), Pt/etched Ni (Figure 10b) and Pt/unpolished Al_2O_3 (Figure 10c) electrodes, respectively. The ($\log P - \log A$) plots of Pt/polished Al_2O_3, Pt/etched Ni and Pt/unpolished Al_2O_3 electrodes gave good linear relations above the threshold area $A_T \approx 3.11\times10^{-13}$, 1.44×10^{-12}, and $7.20\times10^{-13}\ m^2$, respectively. The self-affine fractal dimensions of the film surfaces were determined from Eqs. (5) and (6) by considering all the data points above A_T and the resulting values are listed in Table 2 (2nd column).

Then, ionic diffusion towards self-affine fractal electrode was experimentally investigated by using cyclic voltammetry in a mixture of 30 wt % glycerol and 70 wt % (0.01 M $K_4[Fe(CN)_6]$ + 0.5 M Na_2SO_4) solution. The cyclic voltammograms obtained from each electrode at various scan rates clearly showed one set of well-defined current peaks assigned to be a reversible redox couple as follows:

$$[Fe(CN)_6]^{3-} + e = [Fe(CN)_6]^{4-} \tag{36}$$

Figures 12a, 12b and 12c give on a logarithmic scale the variations of the peak current of the cyclic voltammograms obtained from Pt/polished Al_2O_3, Pt/etched Ni and Pt/unpolished Al_2O_3 electrodes, respectively, as a function of the scan rate. For the very flat Pt/polished Al_2O_3 electrode, the peak current in Figure 12a was linearly proportional to the scan rate with the slope

of 0.500 at all scan rates below the upper threshold scan rate ν_i. However, in the case of rough Pt/etched Ni and Pt/unpolished Al_2O_3 electrodes, the peak currents in Figures 12b and 12c were linearly proportional to the scan rates with the slope of 0.512 and 0.529, respectively, within the scan rate range of the lower threshold scan rate ν_o to the upper threshold scan rate ν_i. It is noted that the temporal inner (short) cutoff τ_i and outer (longer) cutoff τ_o correspond to the upper threshold scan rate ν_i and the lower threshold scan rate ν_o, respectively.

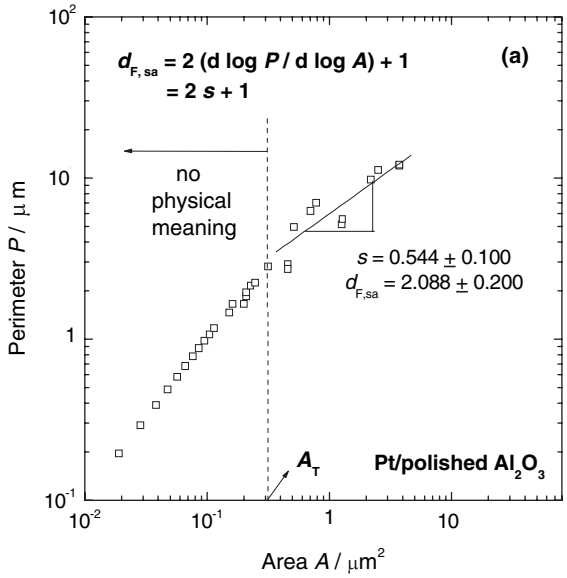
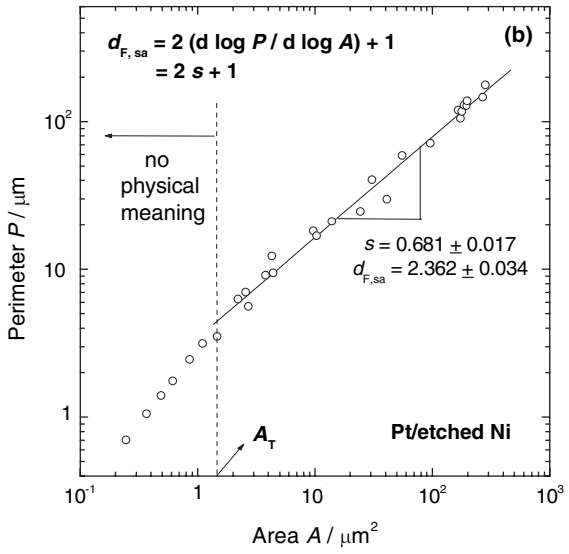


Figure 11. Dependence of the perimeter P on the area A of the lakes on a logarithmic scale obtained at a height corresponding to 40% of the maximum height of the surfaces of (a) Pt/polished AL_2O_3, (b) Pt/etched Ni, and (c) Pt/unpolished Al_2O_3 electrodes. The slope s means (d log P / d log A). Reprinted from J. -Y. Go et al., A study on ionic diffusion towards self-affine fractal electrode by cyclic voltammetry and atomic force microscopy, *J. Electroanal. Chem.*, **549** p. 49, Copyright © 2003, with permission from Elsevier Science.



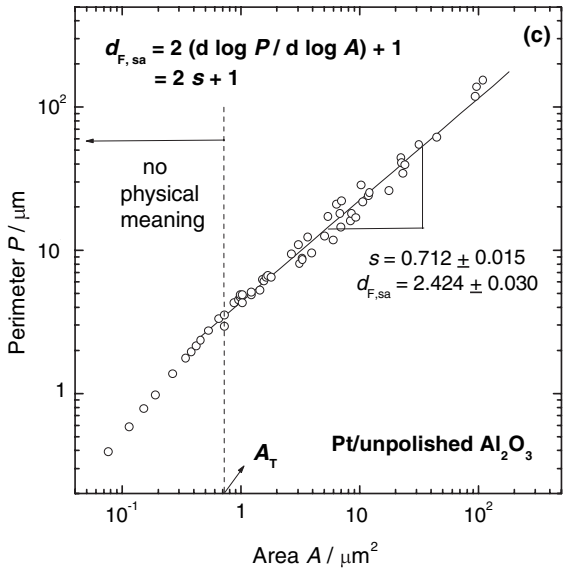


Figure 11. Continuation

Keeping in mind that the *dc* sputter-deposited Pt films have the completely electrochemical-active surface, the fractal dimensions of the rough film surfaces were calculated from Eq. (27) according to the peak-current method by taking the slopes of the (log I_{peak} – log ν) plots within the scan rate range of ν_o to ν_i. The resulting fractal dimensions are listed with the dimension of the flat film surface in Table 2 (3rd column).

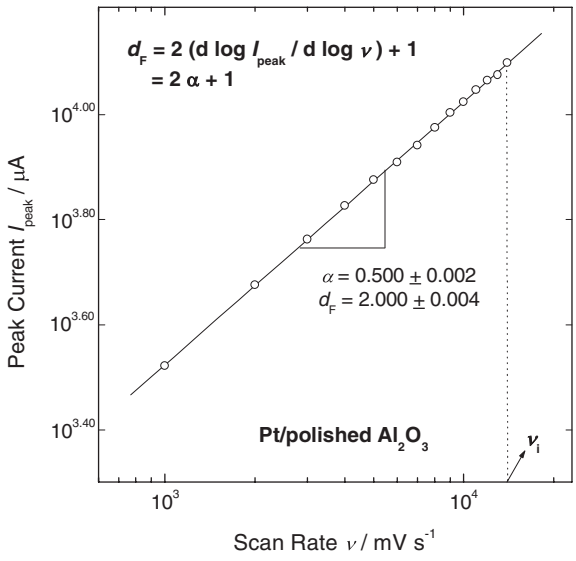


Figure 12. Dependence of the anodic peak current I_{peak} on the potential scan rate ν on a logarithmic scale obtained from the cyclic voltammograms for (a) Pt/polished Al_2O_3, (b) Pt/etched Ni, and (c) Pt/unpolished Al_2O_3 electrodes. The slope α means (d log I_{peak} / d log ν). Here, ν_i and ν_o means the upper and lower threshold scan rate, respectively. Reprinted from J. -Y. Go et al., A study on ionic diffusion towards self-affine fractal electrode by cyclic voltammetry and atomic force microscopy, *J. Electroanal. Chem.* **549,** p. 49, Copyright © 2003, with permission from Elsevier Science.

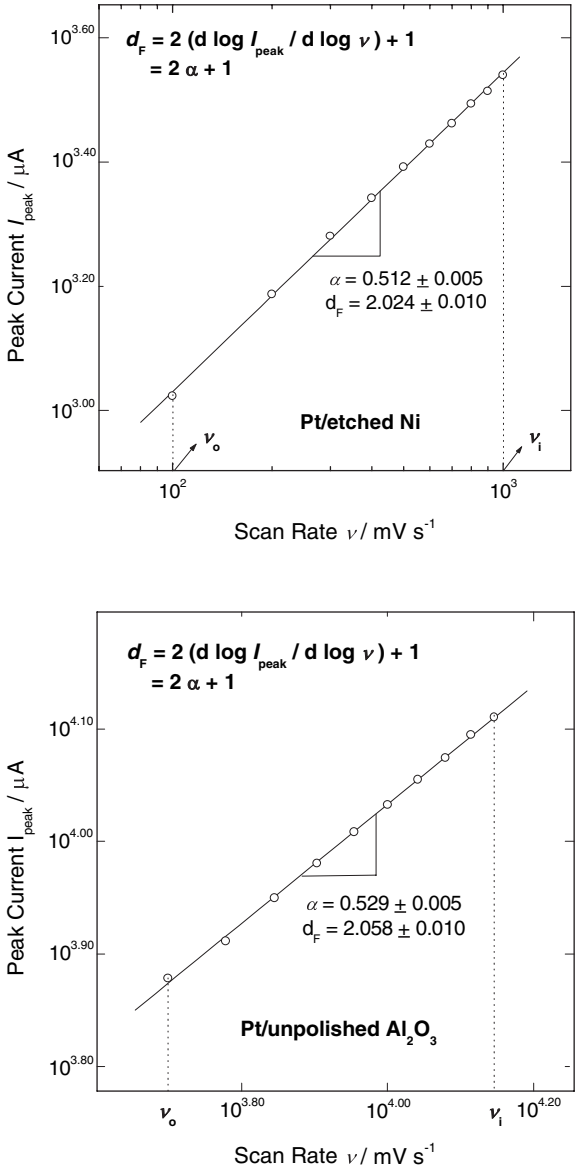


Figure 12. Continuation

Table 2
Fractal Dimensions of Self-Affine Fractal Electrodes Determined by the Perimeter-Area Method (2nd Column), the Peak-Current Method (3rd Column), and the Triangulation Method (4th Column)[a]

specimen \ d_F	$d_{F,sa}$ Determined by the perimeter-area method	d_F Determined by the peak-current method	$d_{F,ss}$ Determined by the triangulation method
Pt/polished Al_2O_3	2.088 ± 0.100	2.000 ± 0.004	2.000 ± 0.000
Pt/etched Ni	2.362 ± 0.034	2.024 ± 0.010	2.020 ± 0.001
Pt/unpolished Al_2O_3	2.424 ± 0.030	2.058 ± 0.010	2.060 ± 0.001

[a] Reprinted from J.-Y. Go et al., A study on ionic diffusion towards self-affine fractal electrode by cyclic voltammetry and atomic force microscopy, *J. Electroanal. Chem.*, **549**, p. 49, Copyright © 2003, with permission from Elsevier Science.

In addition, a length scale range where the rough surface shows the fractal behavior was defined by determining the spatial inner (lower) and outer (upper) cutoffs using Eq. (28).[127,128] The spatial inner (lower) cutoffs for Figures 12b and 12c were determined to be 5.81 and 2.09 μm, respectively, by taking the upper threshold scan rate ν_i and the spatial outer (upper) cutoffs for Figures 12b and 12c were determined to be 19.1 and 3.57 μm, respectively, by taking the lower threshold scan rate ν_o. It is noted that the spatial cutoff ranges for the rough electrodes were in good accordance with the protrusion sizes of the each electrode shown in the AFM images (Figures 10b and 10c). This correlation between the spatial cutoff range and the surface geometry has been found in many other studies.[1,51,54,156]

Now, assuming the self-affine fractal surface can have self-similar scaling property, let us consider self-similar scaling properties of the self-affine fractal Pt film electrodes by the triangulation method. Figures 13a, 13b and 13c give the dependence of the scaled surface area SSA on the projected triangle size TS obtained from Pt/polished Al_2O_3, Pt/etched Ni and Pt/unpolished Al_2O_3 electrodes, respectively. One can find clearly the linear relationship indicating the self-similar scaling properties of the surfaces. The apparent self-similar fractal

dimensions $d_{F,ss}$ estimated from Eq. (4) are listed in Table 2 (4th column).

The values of $d_{F,ss}$ obtained by the triangulation method (4th column in Table 2) were in good accordance with those values determined by the peak-current method (3rd column in Table 2). However, these values are much smaller than the self-affine fractal dimensions determined by the perimeter-area method (2nd column in Table 2). This result is consistent with the theoretical consideration that the self-affine fractal dimension is not always a sufficient condition required for describing ionic diffusion towards self-affine fractal electrode, which is suggested by Shin *et al.*[43]

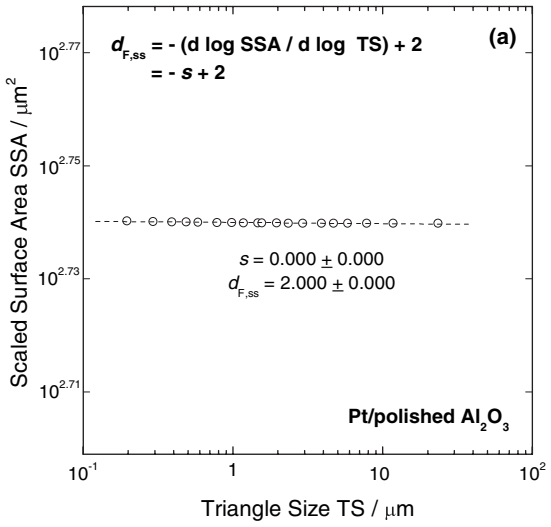
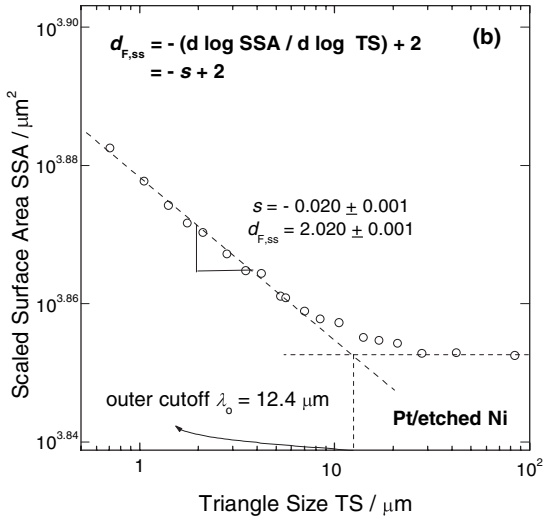


Figure 13. Dependence of the scaled surface area SSA on the projected triangle size TS on a logarithmic scale obtained from the three-dimensional AFM images of (a) Pt/polished Al_2O_3, (b) Pt/etched Ni, and (c) Pt/unpolished Al_2O_3 electrodes. The slope *s* means (d log SSA / d log TS). Reprinted from J.-Y. Go *et al.*, A study on ionic diffusion towards self-affine fractal electrode by cyclic voltammetry and atomic force microscopy, *J. Electroanal. Chem.*, **549**, p. 49, Copyright © 2003, with permission from Elsevier Science.



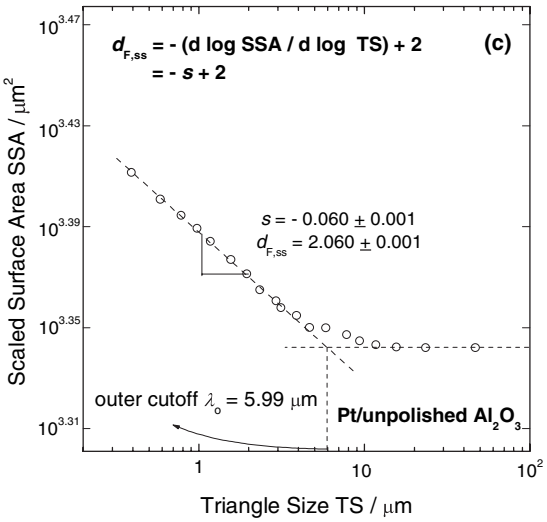


Figure 13. Continuation

Bearing in mind that diffusing ions move randomly in all directions, it is reasonable to say that the diffusing ions sense self-similar scaling property of the electrode surface irrespective of whether the fractal surface has self-similar scaling property or self-affine scaling property. Therefore, it is experimentally justified that the fractal dimension of the self-affine fractal surface determined by using the diffusion-limited electrochemical technique represents the apparent self-similar fractal dimension.[43]

In summary, from the above theoretical and experimental results, it is concluded that ionic diffusion towards self-affine fractal electrode should be described in terms of the apparent self-similar fractal dimension rather than the self-affine fractal dimension. In addition, the triangulation method is one of the most effective methods to characterize the self-similar scaling property of the self-affine fractal electrode.

V. APPLICATION OF THE FRACTAL GEOMETRY IN ELECTROCHEMICAL SYSTEM

1. Corroded Surface

Metallic corrosion is an important field of electrochemistry. Electrodissolution occurring on metal surfaces causes various surface morphologies. The quantitative characterization of the corroded surface using fractal geometry is of great interest since the structure of the corrosion front is a significant factor affecting the transport of ions and corrosion products during corrosion process. To develop a better understanding of the relation between the irregular corrosion front and the physical and chemical corrosion mechanisms, many researchers have devoted their effort to the fractal characterization of the corrosion front.

In earlier studies, the possibilities of the fractal description of the irregular corrosion front were investigated by the image analysis of the digitized micrographs and photographs. Costa et al.[47] studied the shapes of the pits formed on the commercial stainless steel rods, SS 316L, in synthetic ocean water. The self-affine fractal dimension of the pit profiles was evaluated from the analysis of the digitized SEM micrographs using the structured

walk method.[157,158] There were two fractal dimensions for the rugged structure of the pit on shorter length scales ($d_{F,sa}$ = 1.17) and for the texture of the pit on longer length scales ($d_{F,sa}$ = 1.45). They explained this crossover of the fractality in terms of the physico-chemical characteristics of the pitting corrosion.

Jøssang and Feder[48] demonstrated that the corrosion front of a steel pipe subjected to CO_2 pitting corrosion is a self-affine fractal, and that it does not make sense to analyze the self-affine structure as if it were a self-similar fractal. They three-dimensionally reconstructed the corrosion front by digitizing photographs of a sequential grinding of the surface, and developed the image analysis method using height-difference correlation functions obtained from the digitized photographs. The Hurst exponents H in the range of 0.8–0.9 were measured, and it varied with the position of the cut, indicating an inhomogeneous corrosion front.

In order to overcome the difficulty in the observation of the corrosion front, Holten et al.[49] used the two-dimensional cell in which corrosion can take place along only one exposed edge of a metal for the study on the aluminum corrosion in a 1 M NaCl + 0.01 M NaOH (pH = 12) electrolyte. The value of H for the aluminum corrosion front was determined to be in the range of 0.6 – 0.7 consistently using the three different methods: the width-horizontal length method, the height-difference correlation function method, and the power spectrum method. However, the box-counting method did not give reasonable values of H in their work. They extended their approach to zinc and copper, and similar results were obtained.

From the above works, it is confirmed that the irregular surface made by the electrodissolution of the metal can be quantitatively characterized by using self-affine fractal geometry. Based upon this, the pit morphology was extensively investigated by many researchers using a two-dimensional cell in relation to the pit mechanisms. Balázs and Gouyet[50] studied the influence of Cl^- and Fe^{+3} ions on the morphology of pits grown in aluminum thin films in various 1 mM to 1 M NaCl + 0.2 to 10 mM $Fe_2(SO_4)_3$ + 5 mM sulfuric acid solutions. At low Cl^- and Fe^{3+} content (10 mM Cl^- and 0.2 mM Fe^{3+}), the shape of pit was similar to percolation clusters in the vicinity of the percolation threshold. The fractal

dimension of this fractal-like pit was determined to be 1.91 ± 0.01 by the analysis of the digitized images using the density-density correlation function and this value is close to 1.89, which is just the fractal dimension of percolation clusters at percolation threshold in agreement with the suggestion of Sapoval *et al*. for an invasion-like corrosion pattern in the plane.[159]

Recently, Park and Pyun[51-53] have intensively investigated the effect of the solution temperature on the pitting corrosion of Inconel alloy 600 in the Cl^-ion-containing thiosulfate solution. From the SEM observation of the pitted surface of alloy 600 given in Figure 14,[51] pit morphology changed from cylindrical shapes developed at 60 ^oC (Figure 14b) to highly branched shapes formed at 150 ^oC (Figure 14d). Furthermore, corrosion pits nucleated and further grew without any morphological change during the whole pitting process below 200 ^oC. On the other hand, above 200 ^oC, the morphology of the pits changed from highly branched shapes in the early stage to widely grooved shapes in the later stage (Figures 14e, 14f, and 14g).

(a)

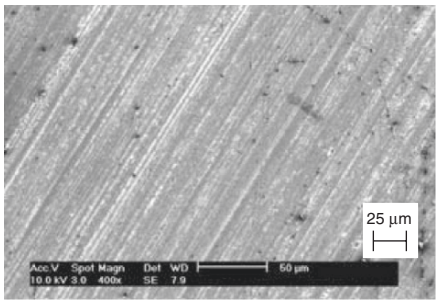


Figure 14. Typical SEM micrographs of pit morphology on the surface of alloy 600 subjected to a constant anodic potential of 0.8 V(Ag/AgCl) for 300s in aqueous 0.1 M $Na_2S_2O_3$ + 0.1 M NaCl solution at (a) 25^o, (b) 60^o, (c) 100^o, (d) 150^o, (e) 200^o, (f) 250^o, and (g) 300^{o}C. Reprinted from J. -J. Park and S. -I. Pyun, Pit formation and growth of alloy 600 in Cl^- ion-containing thiosulphate solution at temperatures 298-573 K using fractal geometry, *Corros. Sci.*, **45**, p. 995, Copyright © 2003, with permission from Elsevier Science.

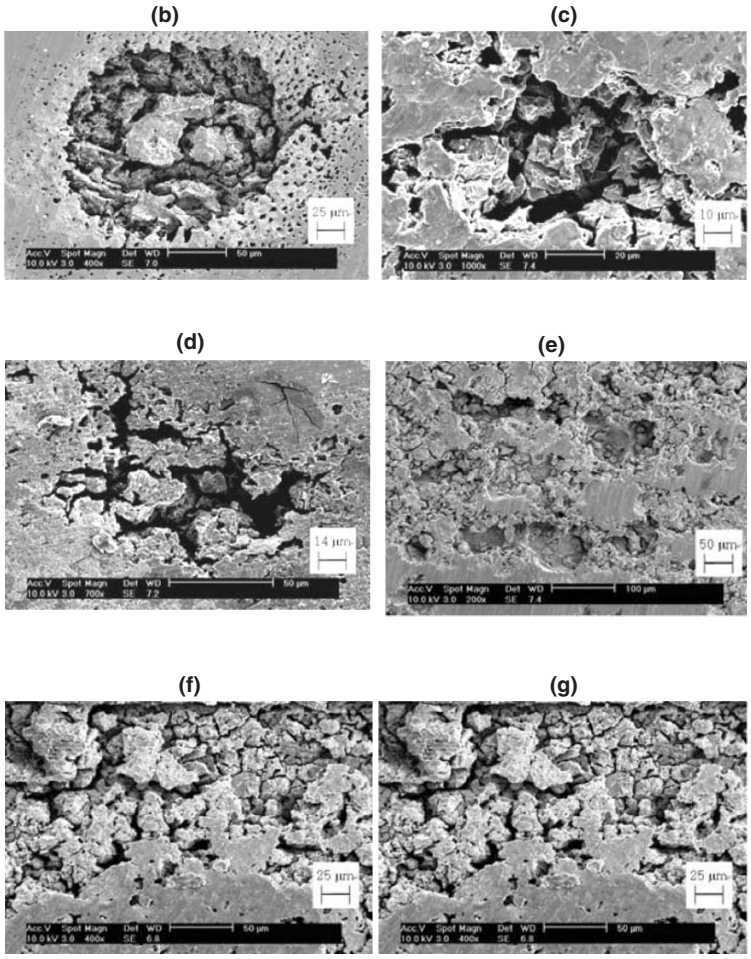


Figure 14. Continuation

In their works,[51-54] the self-similar fractal dimension $d_{F,ss}$ of the two-dimensional distribution of the pits was determined by the analysis of the digitized SEM images using the perimeter-area method. The value of $d_{F,ss}$ increased with increasing solution temperature,[51] and it was inversely proportional to the pit shape parameter and the pit growth rate parameter.[53] Keeping in mind that $d_{F,ss}$ is inversely proportional to the increment of the pit area density, these results can be accounted for in terms of the fact that the increment of the pit area density is more decelerated with rising solution temperature.

In addition, from the analysis of the ac impedance spectra obtained from the previously pitted surfaces under open circuit conditions at room temperature in 0.5 M Na_2SO_4 solution[51] and from the in-situ pitted surfaces in Cl^- ion-containing thiosulfate solution at various applied potentials and solution temperatures,[52] they verified that $d_{F,ss}$ is inversely proportional to the CPE exponent for capacitive charging process (Eq. 30)[51] and the depression parameter for charge transfer process.[52]

The above results concerning the fractal dimension as functions of the solution temperature, the CPE exponent, and the depression parameter are demonstrated again in their other investigation of the effect of the anion addition on pit morphology of Inconel alloy 600 at elevated temperatures.[54]

2. Partially Blocked Active Electrode and Active Islands on Inactive Electrode

Composite electrodes used in the electrochemical processes are often partially active since they are composed of the active powder material and the inactive binder and conductor. The partially blocked active electrode can be characterized by the contiguous fractal with $d_F < 2.0$. In the case of the electrodes composed of the active islands on an inactive support, they are characterized by the non-contiguous fractal with $d_F < 2.0$.[121]

For the contiguous fractal with $d_F < 2.0$, Pajkossy and Nyikos gave the first experimental evidence of the validity of the generalized Cottrell equation.[121] They prepared two kinds of partially active electrodes: a regular fractal pattern with $d_F = (\log 8)/(\log 3) \cong 1.893$ and a Sierpinski gasket[1] with $d_F =$

(log 3)/(log 2) ≅ 1.585. Then they conducted the potentiostatic current transient experiment on two partially active electrodes using the diffusion-limited redox couple of $K_3[Fe(CN)_6]$ and $K_4[Fe(CN)_6]$. For both electrodes, the current transients followed exactly the generalized Cottrell behavior of Eqs. (16) and (24).

Subsequently, using a Sierpinski gasket in a gel electrolyte containing 0.02 M $K_4[Fe(CN)_6]$ + 0.5 M Na_2SO_4 solution, Dassas and Duby[123] demonstrated experimentally that the transition time *vs*. current relationship under galvanostatic control and the peak current *vs*. potential scan rate relationship in a voltammetric experiment as well as the current *vs*. time relationship under potentiostatic control all followed the power laws with the power exponent $-(d_F - 1)/2$ according to the analytical solutions to the generalized diffusion equation involving a fractional derivative operator of Eq. (19).

In practical application, it was reported that the platinum particles dispersed in highly porous carbonized polyacrylonitrile (PAN) microcellular foam used as fuel-cell electrocatalyst[160] have the partially active property. The fractal dimension of the platinum particles was determined to be smaller than 2.0 by using the potentiostatic current transient technique in oxygen-saturated solutions, and it was considered to be a reaction dimension, indicating that not all of the platinum particle surface sites are accessible to the incoming oxygen molecules.

Recently, Lee and Pyun[161] have conducted very interesting research concerning the morphology of surface groups formed and poly-vinylidene fluoride (PVDF)-binder materials dispersed on the graphite composite electrode by using Kelvin probe force microscopy (KFM). They prepared two kinds of electrodes: the as-received SLX50 graphite electrode composed of graphite and PVDF, and the surface-modified SLX50 graphite electrode composed of heat-treated graphite, PVDF, and surface groups on the graphite.

The surface potential profiles of the as-received SLX50 graphite electrode (Figure 15a) showed a smooth surface with a few number of broad peaks, but the surface potential profiles of the surface-modified SLX50 graphite electrode (Figure 15b) revealed a rough surface with many sharp peaks with lower surface potential and road peaks with higher surface potential.

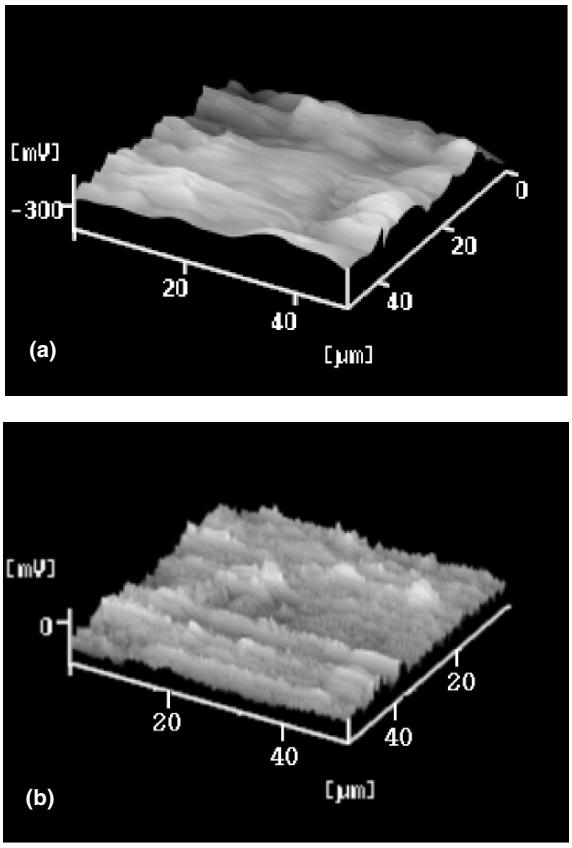


Figure 15. KFM images obtained from the PVDF-bonded composite made from (a) the as-received SFG50 graphite and from (b) the surface-modified SFG50 graphite. Reprinted from S.-B. Lee and S.-I. Pyun, Determination of the morphology of surface groups formed and PVDF-binder materials dispersed on graphite composite electrodes in terms of fractal geometry, *J. Electroanal. Chem.* **556**, p. 75, Copyright © 2003, with permission from Elsevier Science.

Keeping in mind that the surface potential of the PVDF is much higher than that surface potential of the surface groups, they assigned the broad peaks to PVDF and the sharp peaks to surface groups.

In order to determine the fractal dimension of the inactive site distribution, they cross-sectioned the three-dimensional KFM images by the planes with constant surface potential. The self-similar distributions of PVDFs on the as-received SFG50 graphite and the surface-modified SLX50 graphite are given in Figures 16a and 16b, respectively. In addition, the self-similar complex distribution of surface groups and PVDF on the surface-modified SLX50 graphite is shown in Figure 16c. From the analysis of these cross-sectional images by the perimeter-area method, the self-similar fractal dimensions of PVDF on both electrodes and surface groups on the surface-modified SLX50 graphite electrode were determined to be about 1.82 ± 0.02 and 1.70 ± 0.02, respectively.

It is of great interest that they distinguished the fractal dimensions of PVDF and surface groups on the surface-modified SLX50 graphite by using KFM and that they showed the overall fractal dimension determined by the peak-current method is just the average of two individual fractal dimensions determined by the KFM analysis. These results indicate that the surface-modified SLX50 graphite has multifractal geometry.[10]

For the non-contiguous fractal with $d_F < 2.0$ as well as the contiguous one with $d_F < 2.0$, the current response to a potential step obeys the generalized Cottrell equation. Pajkossy and Nyikos theoretically elucidated this fact by the random walk simulation of particles towards a Cantor-like boundary.[121] On the other hand, Strømme et al. experimentally showed the surfaces of the conducting oxide specimens, i.e., In oxide[128] and Sn oxyfluoride,[41,128] are non-contiguous fractal with $d_F < 2.0$. The non-contiguously active sites of specimens were visualized by the hillocks on the specimens observed in the AFM image. Then the fractal dimensions of the hillock distribution on the specimen surfaces were obtained from mass-radius analyses performed on thresholded AFM matrices,[1,63] and these values agreed with that fractal dimensions determined by the peak-current method.

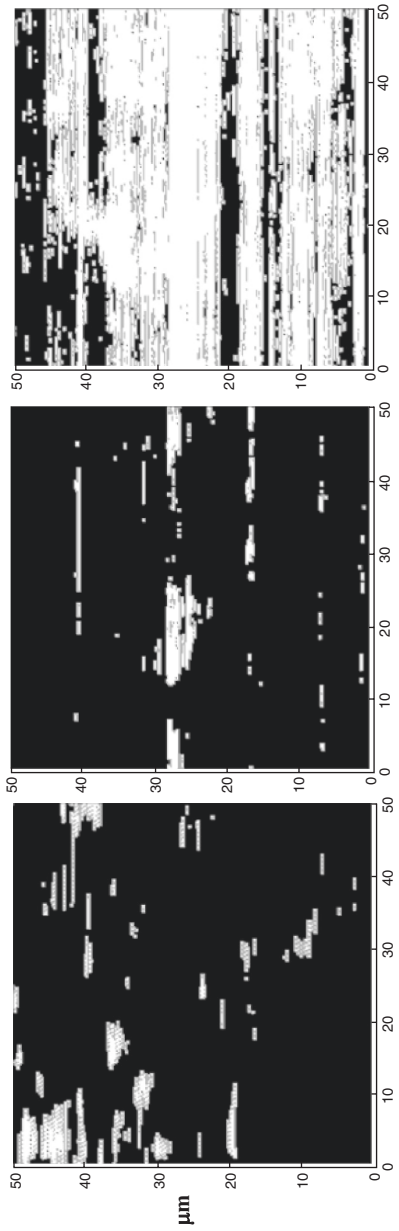


Figure 16. Cross-sectional view of the three-dimensional KFM image (a) for the as-received SFG50 graphite by cross-sectioning with the plane of the height value in 50 % of the maximum height of the surface potential, (b) for the surface-modified SFG50 graphite by cross-sectioning with 80 % of the maximum height of the surface potential, and (c) for the surface-modified SFG50 graphite by cross-sectioning with 50 % of the maximum height of the surface potential. Reprinted from S. -B. Lee and S. -I. Pyun, Determination of the morphology of surface groups formed and PVDF-binder materials dispersed on graphite composite electrodes in terms of fractal geometry, *J. Electroanal. Chem.* **556**, p. 75, Copyright © 2003, with permission from Elsevier Science.

3. Porous Electrode

Porous materials have attracted considerable attention in their application in electrochemistry due to their large surface area. As indicated in Section I, there are two conventional definitions concerning with the fractality of the porous material, i.e., surface fractal and pore fractal.[9-11] The pore fractal dimension represents the pore size distribution irregularity: the larger the value of the pore fractal dimension is, the narrower is the pore size distribution which exhibits a power law behavior. The pore fractal dimensions of 2 and 3 indicate the porous electrode with homogeneous pore size distribution and that electrode composed of the almost same-sized pores, respectively.

On the other hand, the surface fractal dimension characterizes the pore surface irregularity: the larger the value of surface fractal dimension is, the more irregular and the rougher is the pore surface. Since the pore structure is closely related to the electro-active surface area which plays a key role in the increase of capacity in practical viewpoint, many researchers have investigated the microstructure of the pores by using fractal geometry.

Besides the molecular probe method using gas adsorption,[107,162] recently, the TEM image analysis method[163-167] has been applied to evaluate the surface fractal dimension of porous materials. The most attractive fact in this method is that the pores in different size ranges can be extracted from the TEM images which include contributions from many different pore sizes by the inverse fast Fourier transform (FFT) operation by selecting the specific frequency range.[165-167]

The pore structure of pitch-based activated carbon fibers were first investigated by using the TEM image analysis method.[163-166] From the studies on the various activated carbon fibers, it was shown that the fractal dimension of the pores decrease with progressing activation processes for the specimen preparation. The fractal dimension obtained by the TEM image analysis method did not exactly coincide with that fractal dimension evaluated by the molecular probe method using gas adsorption. Keeping in mind that the TEM image gives direct information of the pore structure visually, it was concluded that the fractal

dimension obtained by the TEM image analysis method could more accurately reflect the pore properties of the materials than the molecular probe method using gas adsorption.

Recently, the mesoporous carbons synthesized by the silica imprinting method were investigated by using the TEM image analysis method and the molecular probe method using gas adsorption.[167] In this work, using the TEM image analysis method, two groups of the pores produced from the carbonization of mesophase pitch and from the silica imprinting were distinguished in size and the individual fractal dimension were determined. From the comparison of the overall surface fractal dimension evaluated by the molecular probe method using gas adsorption with the individual fractal dimension, it was recognized that the overall surface fractal dimension is influenced crucially by the individual dimension of the silica-imprinted pore surface.

VI. CONCLUDING REMARKS

The present article summarized the fractal characterization of the rough surfaces and interfaces by using the physical and the electrochemical methods in electrochemistry. In much research, both the physical and the electrochemical methods were used to evaluate the fractal dimension and they are complementary to each other. It should be stressed that the surface fractal dimension must be determined by using the adequate method, according to the inherent scaling properties of the rough surfaces and interfaces.

Regarding the electrochemical method, the generalized forms of the Cottrell relation and the Randles-Sevcik relation were theoretically derived from the analytical solutions to the generalized diffusion equation involving a fractional derivative operator under diffusion-controlled constraints and these are useful in to determining the surface fractal dimension. It is noted that ionic diffusion towards self-affine fractal electrode should be described in terms of the apparent self-similar fractal dimension rather than the self-affine fractal dimension. This means the fractal dimension determined by using the diffusion-limited electrochemical method is the self-similar fractal dimension irrespective of the surface scaling property.

However, in the case of the non-diffusion-limited electrochemical processes, e.g., cell-impedance-controlled lithium transport through transition metal oxides and carbonaceous materials[168] and hydrogen transport through hydride-forming metals and alloys,[169-171] the generalized relations derived under the diffusion-controlled constraint are not valid any more. Therefore, systematic investigation of diffusion towards fractal interface under the non-diffusion-limited constraint has been necessary for the wide application of the electrochemical method to fractal characterization. Since it is very hard to obtain the analytical solutions to the generalized diffusion equation under the non-diffusion-controlled constraint, one of the most useful approaches to diffusion towards fractal interface under the non-diffusion-limited constraint may be the numerical analysis of the generalized diffusion equation.

ACKNOWLEDGEMENTS

The receipt of a research grant under the program 'National R&D Project for Nano Science and Technology', funded by Ministry of Science and Technology, Republic of Korea, is gratefully acknowledged. Furthermore, this work was partly supported by the Brain Korea 21 project.

NOTATION

a	Factor in Weierstrass function
a_0	Diameter of adsorbed molecule
A	Area
A_c	Effective cross-section area
A_{ex}	Exposed surface area of the working electrode
A_{macr}	Macroscopic area which is defined beyond the fractal limits
A_{micr}	Microscopic area which is defined within the fractal limits

A_T	Threshold area above which the linear relationship between the logarithm of perimeter and the logarithm of area appears
α	Power exponent in the generalized Cottrell equation
α_{CPE}	CPE exponent
AFM	Atomic force microscope
b	Factor in Weierstrass function
B	CPE coefficient
β	Proportionality constant in the relationship between the area and the perimeter
B.C.	Boundary condition
c_{ox}^{b}	Bulk concentrations
c_{ox}^{s}	Surface concentrations
C	Power law exponent which is dependent on the surface fractal dimension determined by the single-probe method using gas adsorption
C_e	Electrode capacitance per unit area
χ	Dimensionless parameter
χ_{max}	Maximum value of χ
CPE	Constant phase element
d_E	Topological dimension of the Euclidean space where the set is embedded
d_F	Fractal dimension
$d_{F,B}$	Box dimension
$d_{F,D}$	Divider dimension
$d_{F,MP}$	Surface fractal dimension determined by the multiprobe method using gas adsorption
$d_{F,sa}$	Self-affine fractal dimension
$d_{F,ss}$	Self-similar fractal dimension
$d_{F,ss}^{L}$	Self-similar fractal dimension of the two-dimensional self-similar lakes
$d_{F,S}$	Similarity dimension

$d_{\mathrm{F,SP}}$	Surface fractal dimension determined by the single-probe method using gas adsorption
D	Random diffusion coefficient
$\widetilde{D}_{\mathrm{Ox}}$	Chemical diffusivity of the oxidized species
$\widetilde{D}^{*}_{\mathrm{Ox}}$	Fractional diffusivity
$\widetilde{D}_{\mathrm{Red}}$	Chemical diffusivity of the reduced species
$\partial^{\nu} / \partial t^{\nu}$	Riemann-Liouville mathematical operator of fractional derivative
E	Electrode potential
$E_{1/2}$	Half-wave potential
E_{i}	Initial electrode potential
ε	Threshold for the dominant forces between van der Waals forces and the liquid/gas surface tension forces in the multilayer gas adsorption
F	Faraday constant
f	Frequency
$f_{\mathrm{WS}}(x)$	Weierstrass function
FFT	Fast Fourier transform
FHH	Frenkel, Halsey, and Hill
γ	Dimensionless geometrical parameter
$\Gamma(y)$	Gamma function of y
$h(x)$	Affine function
H	Hurst exponent
η	Morphological amplitude
I	Current
I_{peak}	Peak current
I.C.	Initial condition
j	$\sqrt{-1}$
K	Constant related to the fractal dimension
KFM	Kelvin probe force microscopy
L	Length
λ_{i}	Spatial inner (lower) cutoff
λ_{o}	Spatial outer (upper) cutoff

$\Delta\lambda$	Diffusion layer length
MP	Multiprobe
n	Number of adsorbed molecule layers
N	Number of objects
N_{mono}	Number of adsorbed molecules to form a monolayer
ν	Potential scan rate
ν_{i}	Upper threshold scan rate
ν_{o}	Lower threshold scan rate
Ox	Oxidized species
p	Adsorption equilibrium pressure of the gas
p^0	Saturation pressure of the gas
P	Perimeter
PAN	Polyacrylonitrile
PVDF	Poly-vinylidene fluoride
r	Scaling ratio
$r(S)$	New set of points at position $\mathbf{x}' = (rx_1, ..., rx_{\text{E}})$ generated by the similarity transformation of the set S
$\mathbf{r} = (r_1, ..., r_{\text{E}})$	Scaling ratio vector with the different scaling ratios $0 < r_1, ..., r_{\text{E}} < 1$
$\mathbf{r}(S)$	New set of points at position $\mathbf{x}' = (r_1 x_1, ..., r_{\text{E}} x_{\text{E}})$ generated by the affine transformation of the set S
r_{m}	Radius of the sphere probe molecule
R	Gas constant
R_{o}	Side length of a square electrode
ρ_{el}	Specific resistivity of the electrolyte
Red	Reduced species
s	Slope of the linear line
S	Set of points embedded in Euclidean E-dimensional space at position $\mathbf{x} = (x_1, ..., x_{\text{E}})$
$S_1, ..., S_{\text{N}}$	Non-overlapping subsets of points, each of which is congruent to the set generated by the similarity or affine transformations

σ_g	Generalized Cottrell coefficient
σ_{rms}	Root mean square (rms) roughness
SEM	Scanning electron microscope
SL	Scaled length
SP	Single-probe
SPM	Scanning probe microscope
SS	Segment size
SS_T	Spatial transition point in the $\log SS$ *vs.* $\log SL$ plot
SSA	Scaled surface area
STM	Scanning tunneling microscope
t	Time
t_{ad}	Actual thickness of adsorbed molecule layers
$t_{ad,max}$	Maximum thickness of adsorbed molecule layers
$t_{ad,min}$	Minimum thickness of adsorbed molecule layers
T	Temperature (K)
τ_i	Temporal inner (shorter) cutoff
τ_o	Temporal outer (longer) cutoff
$\tau_{T,cur}$	Temporal transition point in the current transient
$\tau_{T,tri}$	Temporal transition point corresponding to SS_T
TEM	Transmission electron microscope
TS	Projected triangle size
V	Volume of adsorbed gas molecules
V_{max}	Upper limits of the power-law regimes in the plot of $\ln V$ against $\ln\ln(p^0/p)$
V_{min}	Lower limits of the power-law regimes in the plot of $\ln V$ against $\ln\ln(p^0/p)$
V_{mono}	Volume of monolayer coverage
ω	Angular frequency
z	Number of electrons transferred per the oxidized species in the redox reaction
Z	Impedance

REFERENCES

[1]B. B. Mandelbrot, *The Fractal Geometry of Nature*, W.H. Freeman and Company, New York, 1983.
[2]H. -O. Peitgen and D. Saupe, Ed., *The Science of Fractal Images*, Springer Verlag, New York, 1988.
[3]C. A. Pickover, *Computers, Pattern, Chaos and Beauty*, St. Martin's Press, New York, 1990.
[4]H.-O. Peitgen, H. Jürgens, and D. Saupe, *Fractals for the Classroom: Introduction to Fractals and Chaos*, Springer Verlag, New York, 1992.
[5] R. N. Mantegna and H. E. Stanley, *Nature* **376** (1995) 46.
[6]M. H. R. Stanley, L. A. N. Amaral, S. V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M. A. Salinger, and H. E. Stanley, *Nature* **379** (1996) 804.
[7]H. Takayasu and K. Okuyama, *Fractals* **6** (1998) 67.
[8]H. Takayasu, M. Takayasu, M. P. Okazaki, K. Marumo, and T. Shimizu, in *Paradigms of Complexity*, Ed. by M.M. Novak, World Scientific, 2000, p. 243.
[9]P. Pfeifer and M. Obert, in *The Fractal Approach to Heterogeneous Chemistry*, Ed. by D. Avnir, Wiley, New York, 1989, p. 11.
[10] J. C. Russ, *Fractal Surfaces*, Plenum Press, New York, 1994.
[11]J. -F. Gouyet, M. Rosso, and B. Sapoval, in *Fractals and Disordered Systems*, Ed. by A. Bunde and S. Havlin, Springer-Verlag, New York, 1996, p. 263.
[12]R. Brady and R. C. Ball, *Nature* **309** (1984) 225.
[13]M. Matsushita, M. Sano, Y. Hayakawa, H. Honjo, and Y. Sawada, *Phys. Rev. Lett.* **53** (1984) 286.
[14]D. Grier, E. Ben-Jacob, R. Clarke, and L. M. Sander, *Phys. Rev. Lett.* **56** (1986) 1264.
[15]P. Meakin, *Crit. Rev. Solid State Mater. Sci.* **13** (1987) 147.
[16]R. de Levie, *J. Electroanal. Chem.* **281** (1990) 1.
[17]F. Sagues, F. Mass, M. Villarrasa, and J. Costa, *J. Electroanal. Chem.* **278** (1990) 351.
[18]J.M. Gómez-Rodríguez, A.M. Baró, and R.C. Salvarezza, *J. Vac. Sci. Technol. B* **9** (1991) 495.
[19]P. Ocon, P. Herrasti, L. Vázquez, R. C. Salvarezza, J. M. Vara, and A. J. Arvia, *J. Electroanal. Chem.* **319** (1991) 101.
[20]J. M. Gómez-Rodríguez, A. M. Baró, L. Vázquez, R. C. Salvarezza, J. M. Vara, and A. J. Arvia, *J. Phys. Chem.* **96** (1992) 347.
[21]P. Herrasti, P. Ocón, R. C. Salvarezza, J. M. Vara, L. Viázquez, and A. J. Arvia, *Electrochim. Acta* **37** (1992) 2209.
[22]L. Vázquez, R. C. Salvarezza, P. Ocón, P. Herrasti, J. M. Vara, and A. J. Arvia, *Phys. Rev. E* **49** (1994) 1507.
[23]P.L. Antonucci, R. Barberi, A.S. Aricò, A. Amoddeo, and V. Antonucci, *Mater. Sci. Eng. B* **38** (1996) 9.
[24]W. Huang and D. B. Hibbert, *Physica A* **233** (1996) 888.
[25]T. Kessler, A. Visintin, A. E. Bolzan, G. Andreasen, R. C. Salvarezza, W. E. Triaca, and A. J. Arvia, *Langmuir* **12** (1996) 6587.
[26]F. Gobal, K. Malek, M. G. Mahjani, M. Jafarian, and V. Safarnavadeh, *Synth. Met.* **108** (2000) 15.
[27]B. B. Mandelbrot, D. E. Passoja, and A. J. Paullay, *Nature* **308** (1984) 721.

[28]E. E. Underwood and K. Banerji, *Mat. Sci. Eng.* **80** (1986) 1.
[29]C.S. Pande, L.E. Richards, N. Louat, B.D. Dempsey, and A.J. Schwoeble, *Acta Metall.* **35** (1987) 1633.
[30]Z. G. Wang, D. L. Chen, X. X. Jiang, S. H. Ai, and C. H. Shih, *Scr. Metall.* **22** (1988) 827.
[31]C. W. Lung and S. Z. Zhang, *Physica D* **38** (1989) 242.
[32]A. Imre, T. Pajkossy, and L. Nyikos, *Acta Metall. Mater.* **40** (1992) 1819.
[33]R. F. Voss, R. B. Laibowitz, and E. I. Allessandrini, *Phys. Rev. Lett.* **49** (1982) 1441.
[34]J. M. Gómez-Rodríguez, A. Asenjo, R. C. Salvarezza, and A. M. Baró, *Ultramicroscopy* **42-44** (1992) 1321.
[35]M. V. H. Rao, B. K. Mathur, and K. L. Chopra, *Appl. Phys. Lett.* **65** (1994) 124.
[36]L. Spanos and E. A. Irene, *J. Vac. Sci. Technol. A* **12** (1994) 2646.
[37]C. Douketis, Z. Wang, T. L. Haslett, and M. Moskovits, *Phys. Rev. B* **51** (1995) 11022.
[38]G. W. Mbise, G. A. Niklasson, and C. G. Granqvist, *Solid State Comm.* **97** (1996) 965.
[39]M. C. Salvadori, M. G. Silveira, and M. Cattani, *Phys. Rev. E* **58** (1998) 6814.
[40]M. U. Kleinke, J. Davalos, C. Polo da Fonseca, and A. Gorenstein, *Appl. Phys. Lett.* **74** (1999) 1683.
[41]G. A. Niklasson, D. Rönnow, M. Strømme Mattsson, L. Kullman, H. Nilsson, and A. Roos, *Thin Solid Films* **359** (2000) 203.
[42]L. L. G. Silva, N. G. Ferreira, M. E. R. Dotto, and M. U. Kleinke, *Appl. Surf. Sci.* **181** (2001) 327.
[43]H. -C. Shin, S. -I. Pyun, and J. -Y. Go, *J. Electroanal. Chem.* **531** (2002) 101.
[44]X. Sun, Z. Fu, and Z. Wu, *Mater. Charact.* **48** (2002) 169.
[45]J. -Y. Go, S. -I. Pyun, and Y. -D. Hahn, *J. Electroanal. Chem.* **549** (2003) 49.
[46]J. -Y. Go and S. -I. Pyun, *Electrochim. Acta* **49** (2004) 2551.
[47]J. M. Costa, F. Sagués, and M. Vilarrasa, *Corros. Sci.* **32** (1991) 665.
[48]T. Jøssang and J. Feder, *Phys. Scr.* **T44** (1992) 9.
[49]T. Holten, T. Jøssang, P. Meakin, and J. Feder, *Phys. Rev. E* **50** (1994) 754.
[50]L. Balázs and J. -F. Gouyet, *Physica A* **217** (1995) 319.
[51]J. -J. Park and S. -I. Pyun, *Corros. Sci.* **45** (2003) 995.
[52]J. -J. Park and S. -I. Pyun, *J. Solid State Electrochem.* **7** (2003) 380.
[53]J. -J. Park and S. -I. Pyun, *Corros. Sci.* **46** (2004) 285.
[54]S. -I. Pyun and J. -J. Park, *J. Solid State Electrochem.* **8** (2004) 296.
[55]N. Almqvist, M. Rubel, and E. Franconi, *Mater. Sci. Eng. A* **201** (1995) 277.
[56]N. Almqvist, M. Rubel, P. Wienhold, and S. Fredriksson, *Thin Solid Films* **270** (1995) 426.
[57]N. Almqvist, *Surf. Sci.* **355** (1996) 221.
[58]J. C. Arnault, A. Knoll, E. Smigiel, and A. Cornet, *Appl. Surf. Sci.* **171** (2001) 189.
[59]A. Kirbs, R. Lange, B. Nebe, J. Rychly, P. Müller, and U. Beck, *Mater. Sci. Eng. C* **23** (2003) 413.
[60]A. Kirbs, R. Lange, B. Nebe, R. Rychly, A. Baumann, H. -G. Neumann, and U. Beck, *Mater. Sci. Eng. C* **23** (2003) 425.
[61]J. Feder, *Fractals*, Plenum Press, New York, 1988.

[62]R. F. Voss, in *Dynamics of Fractal Surfaces*, Ed. by F. Family and T. Vicsek, World Scientific, New Jersey, 1991, p. 39.
[63]A. -L. Barabási and H. E. Stanley, *Fractal Concepts in Surface Growth*, Cambridge University Press, Cambridge, 1995.
[64]H. E. Stanley, in *Fractals and Disordered Systems*, Ed. by A. Bunde and S. Havlin, Springer-Verlag, New York, 1996, p. 1.
[65]J.F. Gouyet, *Physics and Fractal Structures*, Springer, New York, 1995.
[66]B. B. Mandelbrot, in *Fractals in Physics*, Ed. by L. Pietronero and E. Tosatti, North-Holland, Amsterdam, 1986, pp. 3, 17, 21.
[67]T. Vicsek, *Fractal Growth Phenomena*, 2nd Ed., World Scientific, Singapore, 1991.
[68]A. Majumdar and B. Bhushan, *J. Tribol.* **112** (1990) 205.
[69]A. Majumdar and C. L. Tien, *Wear*, **136** (1990) 313.
[70]A. Majumdar and B. Bhushan, in *Handbook of Micro/Nano Tribology*, Ed. by B. Bhushan, CRC Press, New York, 1995, p. 109.
[71]T. Silk, Q. Hong, J. Tamm, and R. G. Compton, *Syn. Met.* **93** (1998) 65.
[72]K. C. Clark, *Comput. Geosci.* **12** (1986) 713.
[73]J. Liu and T. Furuno, *Wood Fiber Sci.* **33** (2001) 213.
[74]J. Liu and T. Furuno, *Wood Sci. Technol.* **36** (2002) 385.
[75]C. -H. Kim and S. -I. Pyun, *Electrochim. Acta* **48** (2003) 3455.
[76]B. B. Mandelbrot, *Phys. Scr.* **32** (1985) 257.
[77]R. C. Salavarezza and A. J. Arvia, in *Modern Aspects of Electrochemistry*, No. 28, Ed. by B.E. Conway, J. O'M. Bockris, and R.E. White, Plenum Press, New York, 1995, p. 289.
[78]V. Lakshminarayanan, R. Srinivasan, D. Chu, and S. Gilman, *Surf. Sci.* **392** (1997) 44.
[79]A. Majumdar, B. Bhushan, and C. L. Tien, *Adv. Inf. Storage Syst*. **1** (1991) 231.
[80]P. I. Oden, A. Majumdar, B. Bhushan, A. Padmanabhan, and J. J. Graham, *J. Tribol*. **114** (1992) 666.
[81]P. So, E. Barreto, and B. R. Hunt, *Phys. Rev. E* **60** (1999) 378.
[82]A. K. Bisoi and J. Mishra, *Patt. Recog. Lett.* **22** (2001) 631.
[83]D. Avnir and D. Farin, *J. Chem. Phys*. **79** (1983) 3566.
[84]P. Pfeifer and D. Avnir, *J. Chem. Phys*. **79** (1983) 3558.
[85]P. Pfeifer, D. Avnir, and D. Farin, *Surf. Sci.* **126** (1983) 569.
[86]D. Avnir, D. Farin, and P. Pfeifer, *Nature* **308** (1984) 261.
[87]D. Avnir, D. Farin, and P. Pfeifer, *J. Colloid Interface Sci.* **103** (1985) 112.
[88]D. Farin, S. Pelog, D. Yavin, and D. Avnir, *Langmuir* **1** (1985) 399.
[89]A. Y. Myer, D. Farin, and D. Avnir, *J. Am. Chem. Soc.* **108** (1986) 7897.
[90]D. Avnir, *J. Am. Chem. Soc.* **109** (1987) 2931.
[91]D. Farin and D. Avnir, *J. Phys. Chem.* **91** (1987) 5517.
[92]D. Farin and D. Avnir, *J. Am. Chem. Soc.* **110** (1988) 2039.
[93]D. Avnir and M. Jaroniec, *Langmuir* **5** (1989) 1431.
[94]A. V. Neimark, *Adsorpt. Sci. Technol.* **7** (1990) 210.
[95]D. K. Ludlow and T.P. Moberg, *Anal. Instrum.* **19** (1990) 113.
[96]K. Kaneko, *Langmuir* **7** (1991) 109.
[97]M. Sato, T. Sukegawa, T. Suzuki, S. Hagiwara, and K. Kaneko, *Chem. Phys. Lett.* **181** (1991) 526.
[98]D. Avnir, D. Farin, and P. Pfeifer, *New J. Chem.* **16** (1992) 439.

[99]Y. Lefebvre, S. Lacelle, and C. Jolicoeur, *J. Mater. Res.* **7** (1992) 1888.
[100]W. C. Conner and C. O. Bennett, *J. Chem. Soc., Faraday Trans.* **89** (1993) 4109.
[101]K. S. W. Sing, D. H. Evert, R. A. W. Haul, L. Mosocou, R. A. Pierotti, J. Rouquerol, and T. Siemieniewska, *Pure Appl. Chem.* **57** (1985) 603.
[102]R. F. Cracknell, P. Gordon, and K. E. Gubbins, *J. Phys. Chem.* **97** (1993) 494.
[103]K. Kaneko, R. F. Cracknell, and D. Nichlson, *Langmuir* **10** (1994) 4606.
[104]T. Suzuki, K. Kaneko, N. Setoyama, M. Maddox, and K.E. Gubbins, *Carbon* **34** (1996) 909.
[105]K. Kaneko, T. Suzuki, and K. Kakei, *Langmuir* **5** (1989) 879.
[106]K. Kaneko and C. Ishii, *Colloids Surf.* **67** (1992) 203.
[107]M. Sato, T. Sukegawa, T. Suzuki, and K. Kaneko, *J. Phys. Chem. B* **101** (1997) 1845.
[108]P. Pfeifer, Y. J. Wu, M. W. Cole, and J. Krim, *Phys. Rev. Lett.* **62** (1989) 1997.
[109]P. Pfeifer and M.W. Cole, *New J. Chem.* **14** (1990) 221.
[110]P. Pfeifer, J. Kenntner, and M.W. Cole, in *Fundamentals of Adsorption*, Ed. by A.B. Mersmann and S.E. Scholl, American Institute of Chemical Engineers, New York, 1991, p. 689.
[111]I. M. K. Ismail and P. Pfeifer, *Langmuir* **10** (1994) 1532.
[112]P. Pfeifer and K. -Y Liu, in *Equilibria and Dynamics of Gas Adsorption on Heterogeneous Solid Surfaces*, Ed. by W. Rudziński, W.A. Steele, and G. Zgrablich, Elsevier, New York, 1997, p. 625.
[113]J. Frenkel, *Kinetic Theory of Liquids*, Clarendon, Oxford, 1946.
[114]G. D. Halsey, *J. Chem. Phys.* **16** (1948) 931.
[115]T. L. Hill, *Adv. Catal.* **4** (1952) 211.
[116]M. K. Wu, *Aerosol Sci. Technol.* **25** (1996) 392.
[117]P. Tang, N. Y. K. Chew, H. K. Chan, and J. A. Raper, *Langmuir* **19** (2003) 2632.
[118]R. Larraz, *Chem. Eng. J.* **86** (2002) 309.
[119]A. J. Bard and L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications*, John Wiley and Sons, New York, 1980.
[120]L. Nyikos and T. Pajkossy, *Electrochim. Acta* **31** (1986) 1347.
[121]T. Pajkossy and L. Nyikos, *Electrochim. Acta* **34** (1989) 171.
[122]T. Pajkossy, *J. Electroanal. Chem.* **300** (1991) 1.
[123]Y. Dassas and P. Duby, *J. Electrochem. Soc.* **142** (1995) 4175.
[124]T. Pajkossy and L. Nyikos, *Electrochim. Acta* **34** (1989) 181.
[125]M. Strømme, G. A. Niklasson, and C. G. Granqvist, *Phys. Rev. B* **52** (1995) 14192.
[126]M. Strømme, G. A. Niklasson, and C. G. Granqvist, *Solid State Commun.* **96** (1995) 151.
[127]M. Strømme, G. A. Niklasson, and C. G. Granqvist, *Phys. Rev. B* **54** (1996) 2968.
[128]M. Strømme, G. A. Niklasson, and C. G. Granqvist, *Phys. Rev. B* **54** (1996) 17884.
[129]R. de Levie, *Electrochim. Acta* **9** (1964) 1231.
[130]R. de Levie, *Electrochim. Acta* **10** (1965) 113.
[131]A. Le Mehaute and G. Crepy, *Solid State Ionics* **9&10** (1983) 17.
[132]A. Le Mehaute, *J. Stat. Phys.* **36** (1984) 665.

[133]S. H. Liu, *Phys. Rev. Lett.* **55** (1985) 529.
[134]L. Nyikos and T. Pajkossy, *Electrochim. Acta* **30** (1985) 1533.
[135]T. C. Halsey, *Phys. Rev. A* **35** (1987) 3512.
[136]T. Kaplan, L.J. Gray, and S. H. Liu, *Phys. Rev. B* **35** (1987) 5379.
[137]B. Sapoval, *Solid State Ionics* **23** (1987) 253.
[138]Y.T. Chu, *Solid State Ionics* **26** (1988) 299.
[139]R.M. Hill and L.A. Dissado, *Solid State Ionics* **26** (1988) 295.
[140]W.H. Mulder and J.H. Sluyters, *Electrochim. Acta* **33** (1988) 303.
[141]J. C. Wang, *Electrochim. Acta* **33** (1988) 707.
[142]M. Blunt, *J. Phys. A: Math. Gen.* **22** (1989) 1179.
[143]A. Maritan and F. Toigo, *Electrochim. Acta* **35** (1990) 141.
[144]T. Pajkossy and L. Nyikos, *Phys. Rev. B* **42** (1990) 709.
[145]T. Pajkossy, *J. Electroanal. Chem.* **364** (1994) 111.
[146]Z. Kerner and T. Pajkossy, *Electrochim. Acta* **46** (2000) 207.
[147]C. -H. Kim, S. -I. Pyun, and J. -H. Kim, *Electrochim. Acta* **48** (2003) 3455.
[148]A. P. Borosy, L. Nyikos, and T. Pajkossy, *Electrochim. Acta* **36** (1991) 163.
[149]R. Kant, *Phys. Rev. E* **53** (1996) 5749.
[150]R. Kant, *J. Phys. Chem. B* **101** (1997) 3781.
[151]K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, John Wiley & Sons, New York, 1990.
[152]B. R. Hunt, *Proc. Am. Math. Soc.* **126** (1998) 791.
[153]P. P. Trigueros, F. Mars, J. Claret, F. Sagués, J. Galceran, and J. Puy, *J. Electroanal. Chem.* **348** (1993) 221.
[154]A. Maritan, A. L. Stella, and F. Toigo, *Phys. Rev. B* **40** (1989) 9269.
[155]T. Pajkossy, A. P. Borosy, A. Imre, S. A. Martemyanov, G. Nagy, R. Schiller, and L. Nyikos, *J. Electroanal. Chem.* **366** (1994) 69.
[156]F. Normant and A.V. De Walle, *Cartographica* **33** (1996) 1.
[157]H. Schwarz and H. E. Exner, *Powder Technol.* **27** (1980) 207.
[158]B. H. Kaye, *Powder Technol.* **46** (1986) 245.
[159]B. Sapoval, M. Rosso, and J. -F. Gouyet, in *The Fractal Approach to Heterogeneous Chemistry*, Ed. by D. Avnir, Wiley, New York, 1989, p. 227.
[160]S. Ye, A. K. Vijh, and L. H. Dao, *J. Electrochem. Soc.* **144** (1997) 1734.
[161]S.-B. Lee and S.-I. Pyun, *J. Electroanal. Chem.* **556** (2003) 75.
[162]G.-J. Lee and S.-I. Pyun, submitted to *Carbon*.
[163]K. Oshida, M. Endo, S. L. di Vittorio, and M. S. Dresselhaus, *J. Mater. Res*. **8** (1993) 512.
[164]M. Endo, K. Takeuchi, Y. Sasuda, K. Matsubayashi, K. Oshida, and M. S. Dresselhaus, *Electron. Commun. Jpn., Part 2, Electron*. **77** (1994) 98.
[165]K. Oshida, K. Kogiso, K. Matsubayashi, K. Takeuchi, S. Kobayashi, M. Endo, M. S. Dresselhaus, and G. Dresselhaus, *J. Mater. Res*. **10** (1995) 2507.
[166]M. Endo, T. Furuta, F. Minoura, C. Kim, K. Oshida, G. Dresselhaus, and M. S. Dresselhaus, *Supramol. Chem*. **5** (1998) 261.
[167]S. -I. Pyun and C. -K. Rhee, *Electrochim. Acta* in press.
[168]H. -C. Shin and S. -I. Pyun, in *Modern Aspects of Electrochemistry*, No. 36, Ed. by C. G. Vayenas, B. E. Conway, and R. E. White, Kluwer Academic / Plenum Publishers, New York, 2003, p. 255.
[169]J. -W. Lee, S. -I. Pyun, and S. Filipek, *Electrochim. Acta* **48** (2003) 1603.
[170]J. -W. Lee and S. -I. Pyun, submitted to *Electrochim. Acta*.
[171]S. -J. Lee, S. -I. Pyun, and J. -W. Lee, submitted to *Electrochim. Acta*.

5

Phenomenology and Mechanisms of Electrochemical Treatment (ECT) of Tumors

Ashok K. Vijh

Institut de recherche d'Hydro-Québec,
1800, Blvd. Lionel-Boulet, Varennes, Québec Canada J3X 1S1

I. INTRODUCTION AND ORIGINS

Towards the end of nineteenth century, a revolutionary (at that time) and promising surgical technique was introduced in which low-level direct electrical current was used to destroy tumors. It was also realized at that time that the technique involved *electrolytic* processes rather than purely electric ones. Some of the leading research papers of that era, published in the most prestigious journals, carried titles such as:[1-3] "Further observations on the electrolytic dispersion of tumors"; "Electrolysis in gynecology, with a report on three cases of fibroid tumor successfully treated by the method"; "A supplementary paper to Dr. Martin's, citing two cases of fibroid tumor successfully treated by electrolysis". The superior virtues of this technique for tumors and other surgical problems, e.g., *aneurism* (a permanent, blood-filled dilatation of a blood vessel resulting from disease of

Modern Aspects of Electrochemistry, Number 39, edited by C. Vayenas et al., Springer, New York, 2005

the vessel wall) were also extolled in an Editorial of the British Medical Journal at that time.[4] For example, Martin[2] lists a large variety of tumors, polyps, aneurisms, extra-uterine pregnancies (for destroying the foetus in this case), cysts, hernias, hemorrhoids etc., which were *successfully* destroyed by scores of eminent surgeons (Martin gives the exact references to the work of every surgeon in his paper)[2] of that era. Some dissenting voices, however, pointed out at that time that electrolysis was not always successful for treating these problems and that the procedure was painful and dangerous;[5] this failure was attributed, however, to the use of too strong a current.[6]

Franklin Martin, the pioneer of the electrochemical treatment (ECT) of tumors,[2] became a major force in the development of surgery in America; however, towards the end of his life, he regretted that treatments with electricity had fallen into the hands of quacks and charlatans and therefore into disrepute.[7] As was stated in 1991 by Watson:[7] "Even today this area of research is not favoured by grant-giving bodies and is viewed from the same perspective of disbelief." Also, the lack of easy availability of well-controlled and calibrated power supplies, ammeters and voltmeters—and thus the definition of optimum electrolytic conditions—at the end of nineteenth century, caused electrolytic burns and other failures in some patients so that the technique was never integrated into the routine clinical practice.

Although some preliminary and tentative studies[8,9] appeared in more recent times, the resurgence of modern interest in ECT (also called EChT – Electrochemical Therapy) is really due to the rigorous work of Prof. Björn Nordenström (now Professor Emeritus) of the Karolinska Institute, Stockholm, Sweden, member of the Nobel Assembly for Medicine (1967–86) and President of the Assembly (1985). As Head of Diagnostic Radiology at the Karolinska Institute, he became interested in streaks, spikes and coronas that he saw in X-ray images of lung tumors. He began a scientific investigation into the nature of these observations and concluded that these were electrical in origin and were accompanied by movement of ions and water molecules; in other words, these were electrochemical phenomena. He formulated these observations into theoretical interpretations of the naturally occurring electrochemical processes in the body (under different conditions) by proposing Biologically Closed

Electric Circuit (BCEC) pathways within the body.[10] This BCEC theory is an electrophysiological approach that is based on spontaneously occurring electrical potentials and has produced original insights into such fundamental processes as inflammation, wound healing, organ development and differentiation, and, extra-cellular fluid dynamics: the latter is, of course, the movement of ions and water under an electric field gradient viz. electrochemical phenomena such as electroosmosis, transfer of charged particles (i.e., hydrated ions), electrophoretic movement of macromolecules over short and long distances, and, the development (or annihilation) of pH-gradients within the living tissue.

The streaks, spikes and coronas observed by Nordenström, in the X-ray images, around pulmonary malignancies led him to postulate that some parts of this radiographic pattern resulted from the unexplained effect of tumor masses on the distribution of water in the lung tissue. He demonstrated that electrical potentials do exist within lung masses (e.g., between the tumor and the normal tissue) which create a sort of galvanic effect that could alter extracellular fluid dynamics, i.e., the distribution of water in the lung tissue responsible for the coronas observed in the X-ray images.

The aforementioned electrical potentials and the electrochemical events involved in the extra-cellular fluid dynamics caused by them around the tumor (viz. coronas) led Nordenström to propose the use of externally applied potential differences to cause changes that might lead to the destruction of the tumor, i.e., ECT of cancerous tissues. In 1978, he reported on the treatment of lung metastases in 20 patients by the method of ECT,[10] based on earlier preliminary work on five patients. Since then, he has published very extensive studies on the electrochemical treatment (ECT) of tumors.[12-18]

In addition to these[11,18] and other[19] major Swedish efforts, modern work on ECT on animals and human beings is also being carried out in many other countries, e.g., United Kingdom,[7,20] United States of America,[21,22] Australia,[23,24] Canada[25,26] and other countries,[27] especially, Japan[28] and China.[29]

As regards the practical application of ECT to cancer in human beings, i.e., the clinical work, the most extensive investigations have been carried out in China. Inspired by the

work of Nordenström and coworkers,[10,18] and with the active support of the Chinese government, clinical work was initiated there in 1987. Professor Xin Yu-Ling and coworkers at The China-Japan Friendship Hospital in Beijing were the pioneers in this effort and they organized the post-graduate courses for the application of ECT; up to 1998, more than 2000 physicians were trained in ECT.[30,31] In the latest report of this work,[32] clinical applications of ECT on 11000 cancer patients, over a period of 15 years, were reviewed and will be described in a subsequent Section of this chapter. It should be added that the electrochemical treatment of cancer was approved, following an analysis of extensive experimental and clinical data, by a committee of experts organized by the Ministry of Public Health in China, in 1989. Thus the ECT is now an integral part of clinical practice in China and several thousand physicians have been trained in this technique. In other countries, it appears that ECT is still regarded as *experimental* or unconventional and little training is available. Although ECT has been used on many patients in Europe, Australia and Japan as an effective, safe method to treat many types of cancer, it has been tried in only a few cases in the U.S.A. This is because according to a U.S. Food and Drug Administration (FDS) device exemption regulation, ECT was considered a significant-risk device and involves elaborate procedures to get approval for clinical trials. By American standards, ECT is not a well-established method: the pre-clinical studies are sketchy and lack essential data and reliable control of the clinical trials is missing. Although the feasibility of the method for many cancers is not in dispute, it is felt that the method as presently used has not achieved optimal effectiveness. More systematic fundamental research to address pre-clinical issues is needed before ECT can be used on patients in the United States.

II. PHENOMENOLOGY OF ECT

In a typical tumour ECT, a direct current (d.c.) voltage of 8.5 V is applied between two platinum electrodes inserted 3 cm apart in a cancerous tissue (e.g., liver tumour), causing a flow of 30 mA electrolysis current; this current is made to flow continuously for

69 min giving the passage of total charge of 124 C.[33] More generally, voltages of about 6–10 V with d.c. currents of 40–100 mA* are applied to platinum electrodes embedded in the tumour from 1–3 cm apart, with electrolysis being carried out for 1–2 h with the passage of charge being in the neighbourhood of 100–200 C.[29] The main observations are as follows:

1. Water migrates from the anode to the cathode, causing dryness near the anode and oedema in the tissue surrounding the cathode.
2. The anodic *site* in the tissue becomes strongly acidic and the cathodic site strongly alkaline.
3. Higher concentrations of Na+ and K+ in the tissue around the cathode (which has also become strongly alkaline) are observed; not much change in the concentration of multivalent cations such as Ca^2+, Mg^2+, Mn^2+, Cu^{2+} and Al^{3+} occurred during the ECT.
4. The denaturation of protein was caused by the ECT; haemoglobin is converted to acid haemin around the anode and alkaline haemin around the cathode.
5. Chlorine and hydrogen evolution are observed at the anode and cathode, respectively; chlorine causes, not unexpectedly, some bleaching of the local tissue whereas hydrogen produced local cavitation (cf. mechanical damage).
6. The cell metabolism and its existing environment are disturbed severely by the electrochemical treatment, causing the destruction of both normal and tumour cells rapidly and completely.

Although a number of reactions and events, that may contribute to the destruction of tissue, take place during the ECT treatment, it seems that water transport from the anode to cathode is the fundamental event, together with acidity changes associated with the electrode reactions at the anode and the cathode.[25]

We note that before the ECT, the initial pH of the tissue is around 7, i.e. neutral or nearly neutral. When the d.c. power is on during the ECT, there are the usual electrolysis events giving the

* Strictly speaking, current *density* (i.e., mA cm^{-2}) should be specified. In the literature on ECT, only current values are cited although the wire electrodes usually employed have an estimated area of the order of 0.1 to 0.3 cm^2.

evolution of H_2 at the cathode and $O_2 + Cl_2$ (depending on the single electrode potential of the anode, the proportion of O_2 and Cl_2 will change; the anodic potential in turn depends on the parameters of electrolysis) at the anode, as follows

$$\text{Cathode: } 2H_2O + 2e \rightarrow H_2 + 20H^- \tag{1}$$

$$\text{Anode: } 2H_2O \rightarrow O_2 + 4H^+ + 4e \tag{2}$$

and/or

$$\text{Anode: } 2Cl^- \rightarrow Cl_2 + 2e \tag{2a}$$

The evolution of hydrogen at the cathode leads to the build-up of alkalinity; the evolution of oxygen at the anode causes an increase in acidity because the protons generated by Eq. (2) immediately undergo hydration to yield H_3O^+ entities. This is indeed experimentally observed. The chlorine evolution reaction (Eq. 2a) is central and concomitant to the evolution of oxygen.

The electrochemical system involved in ECT may be schematically depicted as in Figure 1.

The total applied voltage, V, in this system is constituted by various components and can be written as

$$V = E^o_{H_2/O_2} + \eta_{\text{anode}} + \eta_{\text{cathode}} + IR + \frac{RT}{F}\, pH \tag{3}$$

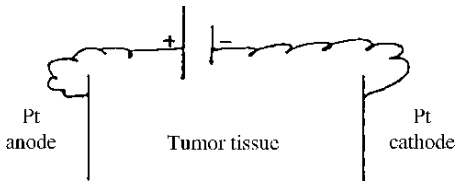


Figure 1. Schematic of the electrochemical system involved in ECT.

where $E^{o}_{H_2/O_2}$, is the reversible potential between the hydrogen and oxygen electrode (= 1.23 V) or between the hydrogen and the chlorine electrode (= 1.36 V); η_{anode} is the anodic overpotential during ECT and its minimum expected value will be ~ 0.3 V, for conditions of detectable oxygen or chlorine evolution; $\eta_{cathode}$ is the cathodic overpotential during ECT and its minimum expected value will be ~ 0.1 V, for conditions of detectable hydrogen evolution; *IR* (*I* is the current, *R* is the resistance and *IR* = *V*) is the resistive drop in the tumour (note: its value during open circuit measurement with a high-input impedance voltmeter will be 0 V; however, during electrolysis involving the passage of high currents, *IR* would become significant, e.g., ~ 1–10 V); (*RT*/*F*) pH is the voltage drop between the anode and the cathode due to the pH gradient developed in the tumour during ECT; at the end of ECT, it can change up to 12 pH units (i.e., 0.708 V); *R* is the gas constant, *F* is the Faraday and *T* is the temperature in degrees Kelvin.

Thus at the beginning of ECT with minimal current (~ 1 mA) and uniform pH (neutral) between the anode and the cathode, the required voltage can be as low as 3–4 V. Towards the end of ECT (high pH gradient, dry anode-tissue interface and thence high IR), especially with a high current (~ 100 mA) and thence high η_{anode} and $\eta_{cathode}$ values, the required voltage could begin approaching 10–15 V.

Although the principle of the ECT method is quite simple, a standarized treatment to make it a reliable, effective method for annihilating cancer tumors is not yet available. Published clinical studies have used different electrode insertion methods and distributions: different electrode placements and configurations have been used in Europe and China. Dosage guidelines (in terms of current, voltage and the charge passed etc.) have been empirically determined to indicate approximate ranges for safety and effectiveness. Excessive dosage will cause pain, burns and slow healing whereas inadequate dosage will result in ineffective treatment. Electrical dosage in tumors varies with factors such as tumors conductivity volume, position and electrode configuration.[34]

Experience[35] has shown that in clinical trials, an appropriate dosage range is 40 to 80 mA current, 8 to 10 V, at 100 C/cm^3 of

charge passed. Note that 100 C/cm^3 of tumor volume is more suitable than the per-centimeter-diameter value since the dose, it is argued,[34] should be related to the volume, and not diameter, of the tumor.

The tumor killing area around an electrode is circular with a radius of approximately 1 cm. Therefore the distance between the positive and negative electrode should be approximately 2 cm. Both electrodes should be inserted into the tumor, or within the peripheries of the tumor in order to avoid damage to the healthy tissue around the tumor. It is preferable to insert the positive electrode (anode) in the centre of the tumor since, in general, tissue necrosis is more pronounced around the anode rather than the cathode. For large tumors, several electrodes may be inserted in order to *cover* the entire tumor for the electrochemical treatment since one anode and one cathode are not effective when they are more than two cms apart.

Based on long clinical experience, the Chinese workers have found that instead of placing an anode in the tumor and a cathode far away from the tumor, *it is better to insert a number of anodes into the centre of the tumor with the same number of cathodes just outside the tumor periphery*.[19,35] Also, ideally, the distance between the anode and the cathode should be between 1 and 2 cm.[37]

As regards safety and toxicity,[34] ECT is a local therapy and no foreign agent is injected into the body — hence one can ignore some toxicity tests such as effects on reproductive functions, teratogenicity, carcinogenicity and mutagenesis. If properly done, ECT provides an attractive alternative to other available means for dealing with cancer, "described by the cynical as cut, burn and poison (surgery, radiation and chemotherapy)," as pointed out to us by Professor John O'M Bockris.

In addition to the main six experimental observations mentioned at the beginning of this Section, some other phenomenological features of the ECT of the cancerous tumors (e.g., grown on the rat skins) have been also noted, as follows:[37]

1. By naked eye observation, one could see colour changes at the anode (yellow) and cathode (white). Also bubbles (H_2) could be seen at the cathode and caseation (i.e., yellowish precipitate, no doubt due to the action of chlorine) at the anode. The anode area appeared *congested* (dry, compact)

and the cathode area was swollen (cf. oedema). Twenty four hours after ECT, the entire tumor shrank and a scar was formed.

2. Light microscope (LM) studies showed that ECT caused distinct changes to the cellular structure of treated tumors. In the anodic area, the nuclei shrank and cytoplasmic structure disappeared.[37] In the cathode area, the cells swelled and the interface among cells blurred. Repeatedly, one observes swelling in the cathode area (oedema) and dehydration around the anode — *this key observation can only be explained in terms of electroosmotically driven water from anode to cathode, to be discussed in details in the Section on Mechanistic Studies.*
3. Transmission electron microscope (TEM) revealed that cellular membranes bursted during ECT and morphological changes corresponding to inhibition of cell proliferation were visible.[37]
4. Both tumor and normal tissues become necrotic around the electrodes after ECT, i.e., the technique lacks specificity in this aspect. For this reason, all electrodes should be inserted into the tumor tissue,[37] or very near the periphery of the tumor.
5. Since the necrosis of the tumor tissue takes place at both electrodes, no fundamental polarity effect is observed.[38]
6. It has been shown that hyperthermia (e.g., by resistive heating of the electrode) plays no role in the tumor destruction.[38]
7. No effect of the *nature* of metallic electrodes was found on the tumor destruction,[7,38,39] e.g., silver and tungsten electrodes gave results similar to platinum:[7] this also means that a slight anodic dissolution of platinum that might occur during the ECT experiment is not responsible for the necrosis of the tumor.[7]
8. It has been observed[40] that, ECT works even if both electrodes are outside the tumor as long as the electrochemical current passes through the tumor:[40] of course, ECT will also destroy the non-cancerous tissue in which the electrodes are embedded. This work of Miklavčič and coworkers also confirms the lack of any role of the nature of electrodes (e.g., Pt, Ir, Au, Ti) or

hyperthermia in the destruction of the tumor tissue, mentioned above from the work of other investigators.[7,38,39]

9. Many workers report that fast-growing regions of organisms, i.e., tumors are electrically negative with respective to the surrounding normal tissue.[9,40,41,42] Consistent with this observation is the fact that crystalline nickel sulphide (α NiS) and cobalt sulphide (CoS_2) which have negatively changed surfaces can cause greater cell transformation and cellular toxicity than the respective amorphous metal sulphide particles which are positively charged.[44]

III. MECHANISTIC ASPECTS

1. General

The mechanisms of the necrosis of the cancer tissue by electrochemical treatment (ECT) are complex and not fully understood although the nature of several factors involved has been indicated. Nordenström pointed out the importance of electroosmosis, electrophoresis, electrode reactions, pH changes and the general drastic change in the microenvironment of the cancer tissue[10,18] during ECT; this and related work has been reviewed by Nilsson and coworkers.[19]

In order to explain all the salient features of the key experimental results on ECT (viz. listed as 1. to 6. at the beginning of Section II, Phenomenology of ECT), Vijh[25] proposed a detailed electrochemical mechanism in which electroosmosis of the tissue (and thence water movement from anode to cathode) and electrode reactions (thence necrosis of the tissue, pH changes etc.) play the dominant roles. In particular, he presented a model and some quantitative considerations that delineate Nordenström's idea of electroosmosis through the narrow interstitial channels lined with fixed charges as the mechanism of the electrochemical destruction of the tumor tissue.[10] Also he examined the role of electrode reactions and other events as possible contributory factors, as follows[25] in Section III.2.

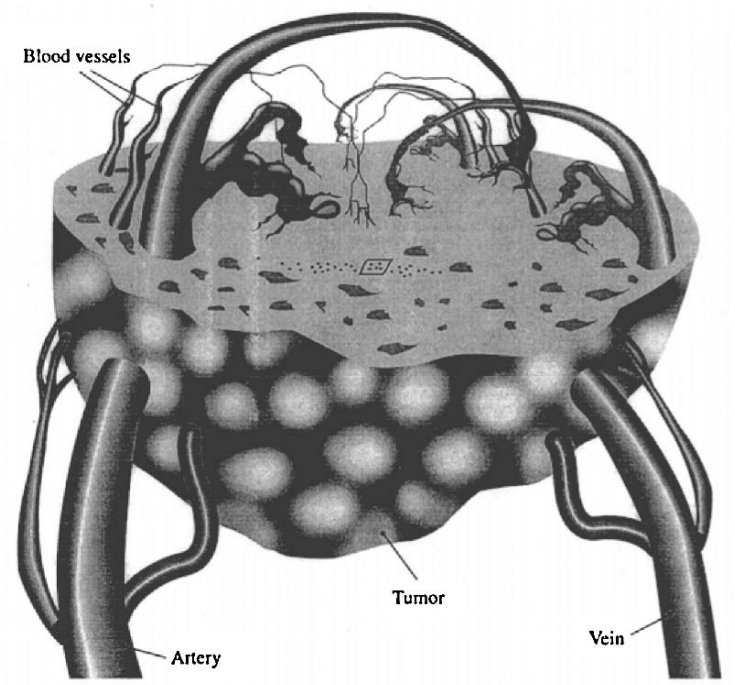


Figure 2. Idealized solid tumor, after Jain.[45] It has been partly cut away to reveal the complex network of blood vessels.

2. Electrochemical Treatment (ECT) as a Case of Electroosmotic Dewatering (EOD) of the Cancer Tissue and Related Effects

A schematic of a solid tumor is presented in Figure 2 and an exploded detail of the small portion of Figure 2 is shown in Figure 3, after Jain.[45] Note the tumor cells (or colonies of tumor cells) floating in the extra-cellular medium called the interstitium (Figure 3). Now we observe that the net charge on the cells of mammals is negative:[10,46] this negative surface charge must be balanced by the positive charge in the extra-cellular fluid at the cell membrane-interstitium interface to form a so-called electrochemical double layer[46] as arises from the electrostatic requirement of electrical neutrality at the interface. One thus

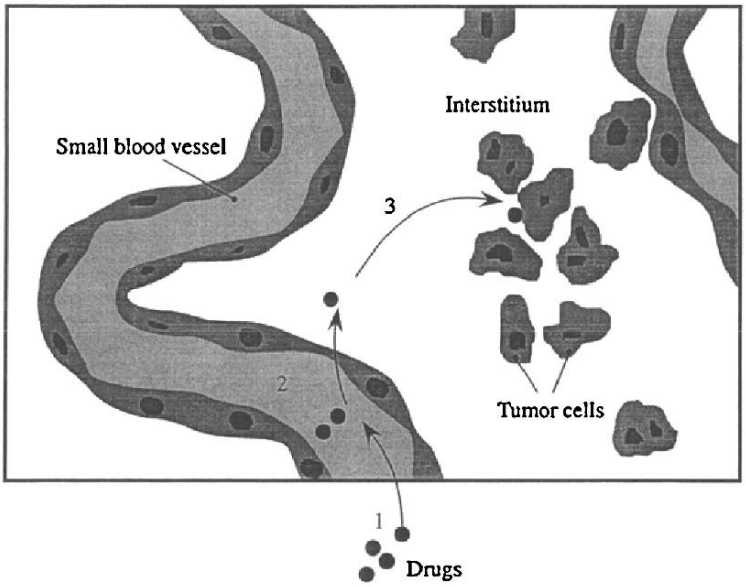


Figure 3. A detailed depiction of a tiny section (in Figure 1) of this tumon showing tumor cells, interstitium, a small blood vessel and the route that a drug i the chemotherapeutic treatment takes. Adapted from Jain.[45]

arrives at a schematic description of negatively charged tumor (cancerous or normal) cells (or groups of cells) floating in the extra-cellular fluid creating surface electrical double layers and with, between the cells (or colonies of cells), narrow intercellular fluid channels in the insterstitium (Figure 4). When a DC current is passed between the two electrodes embedded in the tumor, electroosmotic flow of water occurs from the anode to the cathode, as described below.

The basic point here is that ECT causes a net flow of water from the anode to the cathode, causing electroosmotic dewatering (EOD) of the tissue; the experimentally observed flow of water from anodic site (dry) to the cathodic site (oedema) of the tumor tissue can be explained *only* by an electroosmotic mechanism.

The EOD is based on the electrically-induced flow (viz. Electroosmosis) of water trapped between the tumor cells (Figure 4). Such electrically induced flow is possible because of the presence of electrochemical double layer at the cell-

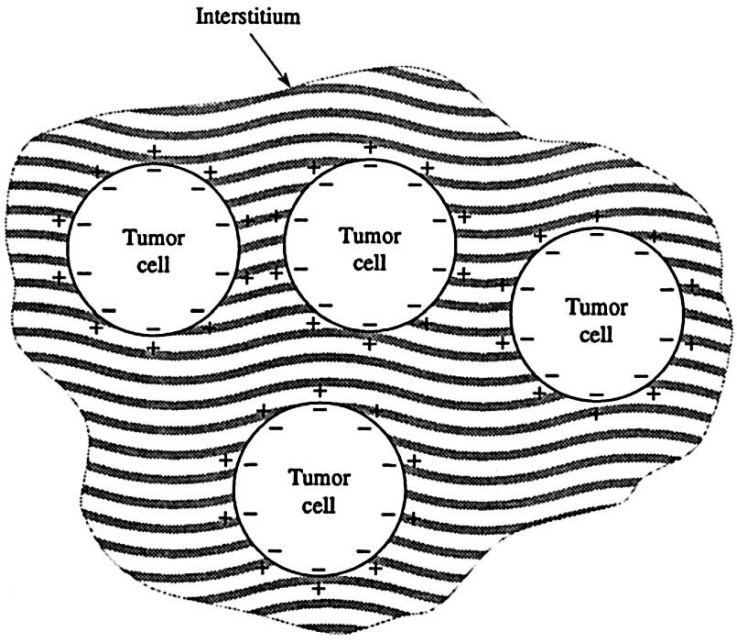


Figure 4. A schematic of tumor cells (or small colonies of tumor cells) with extracellular fluid from the interstitium trapped between them. The electrochemical double layer is indicated for each negatively charged tumor cell with positive ions (in extra-cellular fluid) poised against the negative charge.[25]

interstitium interface; in this double layer (Figure 4), the charges on the cell surface are electrically balanced by the opposite charges in the interstitium, which is actually an electrolyte because of the presence of some salts (e.g., NaCl), hydronium or hydroxyl ions etc. The structure and potential gradients of such a double layer are shown in Figure 5, by analogy with a metal-electrolyte interface.[47]

In this situation, the tumor cells (or a group of cells viz. microtumor) are the immobile phase and the electroosmotic flow causes the water to move as a *plug*, the entire velocity gradient being concentrated at the cell surface in a layer of the same order of thickness as the diffuse double layer (Figure 5). In concentrated solutions, the thickness of the diffuse double is quite small (< 10 Å) whereas in very dilute solutions (as are indeed

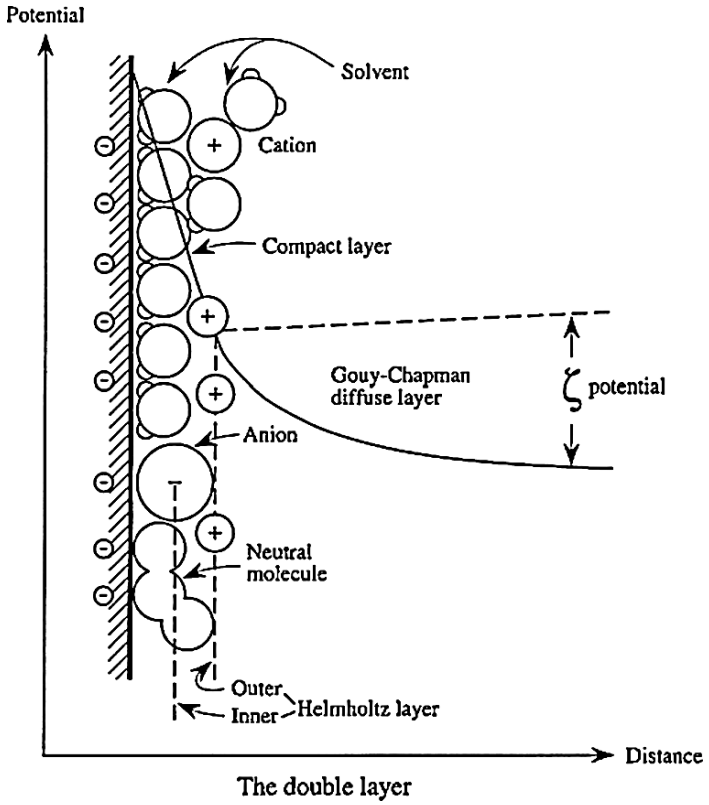


Figure 5. A small portion of the electrochemical double layer at the tumor cell-extracellular fluid (electrolyte) interface is shown to depict the microscopic structure and the potential drops involved, by analogy with the metal-electrolyte interface; taken from Conway.[47]

represented by the water in the interstitium), the diffuse double layer can assume much larger values (~ 10^3–10^4 Å) depending on the concentration of ions in the water. The electroosmotic movement of water between the tumor cells is exactly similar to the electroosmotic flow in a capillary pore (Figure 6): the thin layer of charged fluid (i.e., water containing some ions) next to the cell wall moves like a *single ion* (hence the analogy of a *plug*), under the action of the electric field in a direction parallel to it.[48]

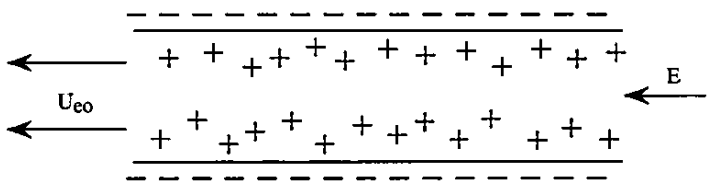


Figure 6. A schematic of the electroosmotic flow of the medium (e.g., an electrolyte) in a capillary caused by the flow of counter ions as a *plug*, under the influence of the applied electric field, E : U_{eo} is the convective liquid velocity from electroosmosis. Adapted from Everett.[48]

The origin of the electrochemical double layer arises from the requirements of charge neutrality in which the surface charge on the cell surface must be balanced against the opposite charge in the water (or any other fluid in a more general case).

Quantitatively, the convective liquid velocity from electroosmosis, U_{eo}, is given by the Helmholtz-Smoluchowski relation:[49, 50]

$$U_{eo}=\frac{\in\zeta E}{\mu} \tag{4}$$

where $\in$ is the permittivity of the extracellular solution; μ is the viscosity of the extracellular solution; ζ, is the so called zeta potential of the cell surface (see Figure 5); and E, is the applied electric strength, defined as the minus of the gradient of the electrical potential (see Figure 6).

The ζ potential is the potential difference between the plane of shear (or slipping plane) and the bulk solution. From Eq. (4), it is clear that for a given situation of water (electrolyte) in the interstitium, the U_{eo} is proportional to the zeta potential and to the applied field strength. Also in a real situation of EOD, it is necessary to use the so called *length-averaged value* of the zeta potential in order to take into account the effect of the axially variable zeta potential on the electroosmotic velocity.

Further, since both the zeta potential and the electric field depend on the ionic concentration and pH of the fluid (the tissue

electrolyte in our case), the variations in electroosmotic velocity can be expected as the dewatering experiment progresses in time.

It should be noted from Eq. (4) that for the EOD to occur, the passage of electric *current* is not required, if *ideally* one could develop high E (V cm^{-1}) without the passage of significant electrical current. However, in practice, the cell-interstitium medium has a given resistance R so that

$$i = \frac{V}{R} = \frac{E}{R} \quad \text{(for tumor thickness = 1 cm)} \tag{5}$$

It is the passage of this *current* that causes the electrochemical reactions at the two electrodes inserted in the tissue for applying the potential gradient (i.e., field) necessary for the electroosmotic flow to occur. The electrode reactions at the anode and cathode induced by the passage of this current result in a number of effects, e.g., changes in pH near the electrodes which give rise to concentration gradients in the bulk of the tumor causing changes in ζ values near the electrodes which result into reduced electroosmotic flow, as the ECT proceeds in time.

The initial pH of the tissue is around 7, *i.e*, neutral or nearly neutral. When the DC power is on during the ECT, there are the usual electrolysis events that cause alkalinity at the cathode and acidity at the anode viz. Eqs. (1), (2) and (2a), and related text in the Section II, Phenomenology of ECT.

It is important to emphasize here that during electroosmosis, the water flow is *always* from anode to cathode, as indicated in the schematic in Figure 7 here. This arises from the electrostatics of the situation in which the water velocity profile follows the direction of the electric field, i.e., from the positive electrode (anode) to the negative electrode (cathode), as depicted in Figure 8.

Thus the foregoing analysis clarifies the two main observations of ECT, namely, the dehydration and acidity at the anode, and, excess hydration and alkalinity at the cathode. It will also explain the relative higher concentrations of Na^+ and K^+ at the cathode since $C\lambda^-$ ions must have migrated to the anode and gotten discharged there, leaving a relative excess of Na^+ and K^+ in the vicinity of the cathode at the end of the ECT.

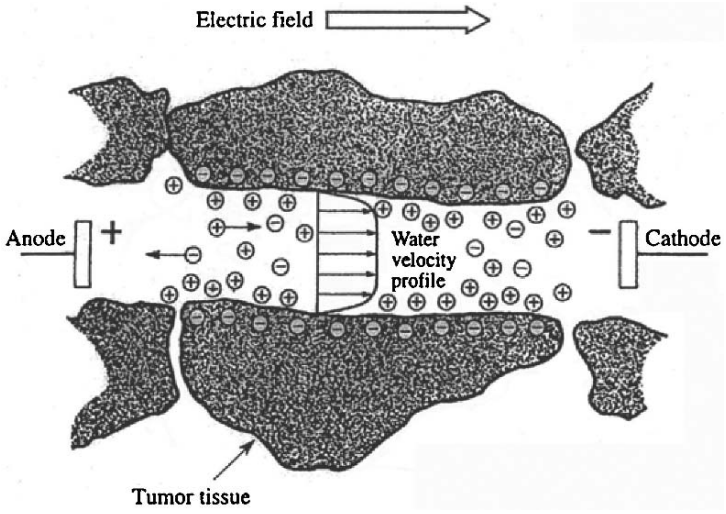


Figure 7. A schematic of the electroosmotic experiment during ECT of a tumor.

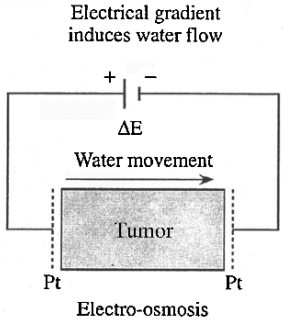


Figure 8. Electroosmosis depicted in the negatively charged tumor tissue wall enclosing an interstitial channel carrying extracellular liquid; the water movement follows the direction of the field, i.e., always from the anode to the cathode.[25]

3. Some Possible Factors in the Necrosis of Cancer Tumors by ECT

The most important factor in tissue death during ECT appears to be dehydration in the anode area and excess hydration (oedema) in the cathode area. This factor is reinforced by high acidity at the anode and excess alkalinity at the cathode. Hyperthermia would appear to be unimportant except perhaps towards the end of ECT experiment when the anode area is quite dry and thence of high electrical resistance and supports enhanced local heating.[25]

Oxidation by $C\lambda_2$ (or chloride radicals adsorbed at the platinum electrode as *Pt*-Cλ), O_2, O_3 (which must be evolved in traces at the high anodic potentials attained in ECT carried out at high current densities) at the anode are no doubt responsible for necrosis in the tissue around the anode; one can even see visible changes such as bleaching and/or other colour changes. Around the cathode, the erosive damage may be caused by cavitation by the profusely evolving hydrogen. Also, hydrogen evolution would expel any oxygen in the tissue around the cathode so that *respiration* by the tumor cells needed for their survival and growth would be inhibited, leading to their necrosis.

The *traumatic* changes brought by ECT would also tend to break and dismantle the hydrogen-bonding network of DNA of the tissue thus suggesting another mechanism for necrosis.[25] The following other possible contributions to the necrosis of the cancer tumor during ECT may also be noted:

1. *Micro-Thromboses:* Platelets and leucocytes, like other cells, are known to carry surplus negative charge, and can be electrophoretically deposited at (or around) the anode[10] owing to the positive electrode potential. These microthromboses in capillaries in or near the treatment site will result in decreased blood flow and may contribute to a local dystrophy of the tissue. A more pronounced version of this effect can be seen sometimes as electrocoagulation and vascular occlusion (i.e., shut, closed or obstructed vasculature) of the tumor tissue. In other words, ECT cuts off the blood supply to the tumor and causes the tumor cell necrosis.
2. *Reactive Oxygenated Species* (*ROS*): Not only can reactive radicals (arising from oxygen or chlorine evolution) be

produced at the anode, but reduction of any oxygen available around the cathode can give rise to H_2O_2 and peroxide damage of the cancer tissue;[26] alternatively, hydrogen evolution at the cathode can cause hypoxia (oxygen deficiency) and tissue death. Since solid tumors are hypoxic (oxygen deficient), cathodic reduction of O_2 to produce H_2O_2 etc. would not be an important reaction.

3. *Other Factors*: ECT no doubt destroys the homeostasis of the cancerous cells by profoundly disturbing their micro-environment and by dismantling their structure. It might also lead to the inhibition of their DNA synthesis. It has been speculated that ECT might release the Tumor Necrosis Factor (TNF) or, somehow super-activate the immune system at the tumor site.

It has been suggested[10,18] that the electroosmotic movement of water towards the cathode that causes oedema at the cathode site leads to increase of *tissue turgor pressure* there:[18] it results in completely suspended circulation in the capillaries in the cathodic field leading thus to tissue death.

4. Approaches to the Possible Improvement of the Electrochemical Treatment (ECT) of Cancer Tumors

Based on considerations of fundamental electrochemistry, ideas were presented in 1999 to improve the application of the ECT technique;[25] to our knowledge, this approach has not been tried experimentally. The argument involved is as follows:

By drawing analogy from the case of electroosmotic dewatering (EOD) of electrolyte trapped in the interstices between fine clay particles (cf. Figure 2 in which tumor cells are replaced by negatively-charged fine clay particles), one notes that electrode reactions at the anode and cathode cause pH changes that reduce the rate of water removal as the ECT proceeds; in other words the initial ζ potential (cf. Eq. 4) assumes lower values with time (due to pH changes) as the ECT proceeds. In order to restore the original ζ potential value, one should gradually lower and then stop the current and short-circuit (after disconnecting the power source) the anode and cathode for a couple of minutes before restarting the ECT; this short-circuiting restores the original ζ

potential and thence the high rates of dewatering observed at the commencement of the process.[49,50] Also, subsequent to the short-circuiting, a *reversal* of the direction of the current for a minute or so also achieves the same goal in a more enhanced manner; this latter would then also have roots in the suggestion of Nordenström[18] in which a reversal of the current in the tissues is achieved. In electrochemical terms, this restores the high ξ potential (Eq. 4 here) and therefore causes the recommencement of EOD at high rates once the normal polarity of the ECT is re-established.

Other possible improvements in ECT are the suggested benefits of combining ECT with other established treatments such as chemotherapy[51] or radiotherapy;[52] both these studies found the combination therapy more effective. Also, immunologically-related compounds such as TNF-α, IFN-α and IL-2 were observed to enhance the effect of ECT when used in combination with the latter.[53,55] The studies have also been covered in a recent synoptic review.[20] This combination therapy was also pioneered, it seems by Nordenström in Sweden;[14,15] he also describes the details of the electrode wire electrodes, cannula and catheter etc. he used in his work.[14]

Another possible avenue for the understanding and improvement of ECT is through mathematical modelling of the ECT, as has been explored by Nilsson and coworkers.[57,58] In one study,[57] they concluded that during ECT, the direct current is not transported homogeneously in the tissue but through thin channels of institial fluid, which is consistent with the detailed electroosmotic theory of ECT[25] described in Section III.2 of this Chapter. In another investigation,[58] Nilsson et al. carried out simulations by modelling tissue as an aqueous solution of sodium chloride containing a bicarbonate buffer system and organic constituents susceptible to reactions with chlorine. They concluded that in ECT at moderately high current densities, it is the spreading of hydrogen ions—and not chlorine molecules—that determines the range of tissue destruction around the anode; this, again, is consistent with the observations of many workers[59] and the facts that H_3O^+ migrates through the tissue much faster than Cl_2, and, large shifts in pH are observed during the ECT.[25] Thus the tissue necrosis by chlorine-based species will be in the

immediate vicinity of the anode whereas tissue destruction by low pH (i.e., H_3O^+) will cover a much larger area around the anode.

5. In-Vitro Studies on the Suppression of Cell Proliferation by the Passage of Electric Current

It is somewhat surprising that only a few investigations are available on the *in-vitro* studies of the suppression of cancer cell proliferation by the passage of electric current; almost all publications deal with the *in-vivo* work on solid tumors.

Lyte and Coworkers[60] examined the effect of low level DC current on some mouse lymphoma cells; within a narrow current range (10–24 μA), some inhibition of the cell proliferation was observed whereas *an enhancement* of cell proliferation was noted outside this current window. The data[60] thus appear somewhat contradictory and, in any case, were not interpreted as due to electrochemical effects although a vague relevance of this work to investigations of Nordenström[10-18] was implied and an electrochemical mechanism thus insinuated.

An *in-vitro* electrochemical treatment (ECT) of KB cells from a human oropharyngeal carcinoma line has been carried out by Yen and coworkers;[61] inhibition of cell growth in culture was observed and found to depend on the magnitude of current and treatment time. Results were interpreted in terms of pH effects caused by ECT.

In another detailed study,[62] the target tumor cells for the DC treatment were mouse mastocytoma P815. The current was passed in a three compartment cell in which the cathode compartment (CC) was separated from the anode compartment (AC) by the insertion of an intermediary chamber (IC) which had neither the anode nor the cathode; the three compartments were *connected* in series via filter-paper bridges (Figure 9). Each of the three compartments contained 2 ml of cell suspension and 2 mA current was passed through platinum electrodes inserted in the AC and CC chambers (Figure 9). The salient results of this study are as follows:

1. In the cathodic chamber, cell lysis and increase in pH was caused by the ECT; however, when the cells were treated in phosphate buffer—a stronger buffering system that does not allow pH variations as a consequence of DC

treatment—neither of these effects were observed. Thus in the cathode chamber ECT causes increase in pH by the production of OH^- (cf. Eq. 1 here) which in turn causes cell lysis.

2. In the anodic compartment, ECT leads to greatly diminished cell viability; however, if instead of ECT, the cells were incubated in a solution of the same pH as was obtained in the anodic compartment after the passage of current, cell viability did *not* diminish. One may conclude that pH is not the main factor in the cell viability in the anodic compartment.
3. When the regular saline solution is replaced by a phosphate buffer in the anode compartment (which means absence of

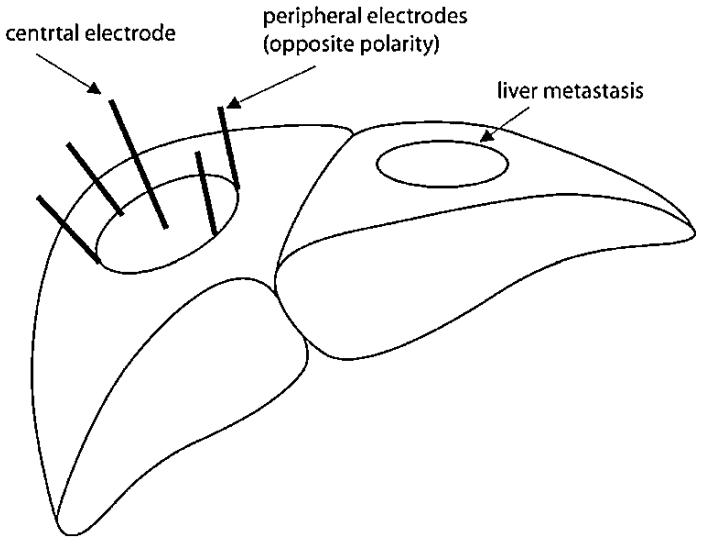


Figure 9. Schematic representation of the acrylic chambers used for treatment of P815 cells with DC. Chambers are connected in series by filter-paper bridges, and fitted with platinum electrodes in their extremities. In this system, cell suspensions can be exposed directly to the anodic reactions (AC) or cathodic reactions (CC) or to electric current without contact with the electrodes, in the intermediary chamber (IC). Internal volume: 3 cm^3. After Veiga et al.[62]

chloride ions) ECT experiment does *not* diminish the cell viability. Hence the annihilation of cancer cells in the anode compartment during ECT is due to chlorine evolution and does not take place owing to pH change.

4. After the ECT, the cells in the intermediary chamber (IC in Figure 9) have the same cell viability as the untreated cells. This means that the effect of ECT arises from the *electrochemical reactions* (cf. anode and cathode compartments) and that its origin does not lie in purely *electrical* effects that may be associated with the passage of current or the presence of voltage drop and thence the electrical field (cf. intermediary chamber).
5. It was observed[62] that ECT induces changes (i.e., in the anodic and cathodic compartments) such as cell membrane damage, alterations in cell shape and chromatin organisation, mitochondrial swelling and condensation, cytoplasmic swelling and matrix rarefication; in other words, profound and destructive changes in the cellular structure and environment, as was also mentioned in Section II, Phenomenology of ECT of solid tumors.
6. An additional[62] major alteration at the cell surface of P815 cells subjected to ECT was disturbance and disruption of carbohydrate expression. Carbohydrate-mediated interactions are involved in the adhesion of cancer cells to endothelial cells, which is correlated with metastasis.[63] These carbohydrate groups help negotiate the cell's relationship to its external environment. Cancer cells speak a different *sugar dialect* than do the normal cells;[64] if ECT were to inhibit specific glycosylation reactions, as indicated,[62] it might be able to alter these pathways and turn off the cancer cell. Hence, one possible way in which ECT can destroy the tumors is by modifying and/or disrupting the carbohydrate groups on the surfaces of cancer cells.

IV. ELECTROCHEMICAL TREATMENT (ECT) OF TUMORS IN ANIMALS

Although the ECT work was initiated on human beings[1-6] in the nineteenth century, the more recent work involves *in-vivo* investigations on a variety of *model* animals at least in the occidental world; in parallel, there is considerable clinical work in China,[30-32] and to a lesser degree, in some other countries.[18,28,34] A survey of some representative ECT studies on animals is presented next.

1. Mice

Humphrey and Seal[9] (U.S.A.) reported in 1959 detailed experiments on the ECT of sarcoma-180 tumors in nearly 500 mice. In a typical experiment, 18 test animals and 18 control animals were used; the ECT was done with 3 ma current, at 6 V for 3 hours a day, for 24 days. They observed that after such a lengthy treatment, the mean volume of test tumors was about 15% of the mean tumor volume of the control group; seven mice showed complete tumor regression as a result of ECT. Total regression means that the tumor has decreased progressively in volume, hardened, dropped off, leaving a new skin surface at the former tumor site.

In some related experiments,[9] they concluded that ECT did not show any promise in the treatment of leukemia in mice induced by intraperitoneal L1210 tumor tissue implant; in other words, ECT is useful for solid tumors only.

The work of Ito and coworkers[65] (Japan) on fibrosarcoma in mice showed that ECT retarded the tumor growth in mice; the best results were obtained when ECT was preceded by a radiation treatment.

Morris et al.[66] (U.S.A.) induced Lewis lung carcinoma (LLC) in C57 black mice and subjected them to ECT: constant current of 20 mA (which gave rise to 8–10 V as the voltage across the electrodes) was applied for 15 minutes. Animals treated with ECT survived longer and exhibited a small primary tumor burden at death as compared to the control group, although no complete regression of the tumor was observed. It would appear that a

treatment time of 15 minutes was probably too short to kill the tumor.

Griffin and coworkers (U.K.) reported work on ECT of mammary carcinoma in mice.[67] Tumor destruction was found to be significantly greater, for a given charge, when anode was implanted in the tumor with cathode outside it. The results demonstrated a linear relationship between the volume of tumor regression induced and the quantity of electric charge passed.

A very active group in Slovenia involving Miklavčič, Serša and coworkers have done extensive work on the ECT of tumors in mice.[68-72] The effect of direct current (0.2–1.8 mA) on fibrosarcoma and melanoma tumors in mice was examines for 30–90 minutes, either in single dose or repetitive treatments. Several electrode materials were investigated, e.g., Pt, Pt-Ir, Au, Ag, Ti and stainless steel. A number of electrode configurations were also explored, e.g., one electrode placed in the tumor and the counter electrode in the healthy tissue, *or*, so called the *field configuration* where both electrodes were placed in the healthy tissue on opposite sides of the tumor. A decrease in the tumor volume was noted with all electrode materials and configurations. The tumor growth retardation was proportional to the electrical charge passed. The best ECT effect was achieved by using a multi-array electrode.[69] The other interesting work of this group, mentioned in a previous Section,[51-56] uses ECT in combination with bleomycin or immunomodulators such as tumor necrosis factor (TNF-α), interferon-α (IFN-α), interleukin-2 (IL-2) etc., with much beneficial effects.

In the work of Cabrales and coworkers[73] (Cuba), sixty male BALBC mice were used and Ehrlich tumors were implanted in them. When the tumor volumes reached 850 mm^3, four platinum electrodes were inserted into each tumor: ECT was carried out at 4 mA for 21 minutes, total charge thus passed through each tumor was around 5 coulombs. Tumor volume decrease and necrosis percentage increase were significant in mice treated with ECT, in comparison with the control group without ECT.

It is clear that in these studies on the ECT of solid tumors in mice, an appreciable reduction in the tumor volume or its complete necrosis is generally observed.

2. Hamsters

In their work[74] on the ECT of melanoma tumors in hamsters, Schauble et al. (U.S.A.) implanted a 1.5 V battery as the power source. A stainless steel point electrode was placed in the tumor, either as anode or as cathode and a counter electrode made of plane wire mesh was applied to the skin of the chest. Three different current levels were used: 3 mA; 0.5 mA; 1 μA. The two higher currents produced necrosis of the tumor at the dose leve of 1h/day for 4 consecutive days. A more pronounced effect was observed with anode in the tumor and cathode outside it.

David and coworkers (Canada) investigated the effect of ECT on amelanotic melanoma (T1-4) solid tumors in hamsters.[75] The treatment consisted of a passage of currents from 0.1 to 2.4 mA, for 1h/day for five consecutive days. Treated tumors decreased in mass (as a percentage of controls) to 89% at 0.1 mA and to 2% at 2.4 mA. Similar results were obtained with both stainless steel and platinum—30% iridium electrodes. It was observed that even at the higher current level, a thin rim of viable cells still remained around the periphery after the treatment: it would thus appear that the treatment did not use high enough currents for sufficiently long periods and the electrode configuration probably did not span the entire tumor during ETC.

3. Rats

Gardner et al.[75] (U.S.A.) injected ascetic Walker sarcoma cells into male albino Wistar rats and thus induced tumors in their liver. These tumors were than subjected to ECT by using an implanted 1.5 V mercury battery as the power source; platinum electrodes were used. Although no mention was made of the level of current passed, it must be quite low based on the battery voltage and the resistors used in their circuit. In any case, they observed that a lot of cells clumped around the anode—they interpreted this as the attachment of negatively charged cancer cells to the positive electrode. Obviously, the ECT treatment was not strong and long enough to cause necrosis or tumor regression etc.

Habal[77] (U.S.A.) used small direct current (0.4–0.6 μA, 10–24 days) drawn from an implanted power source in his ECT of hepatoma tumors in rats. The power sources (1.5 V) were

implanted after different times following the tumor implants in rats; the tumor growth retardation only occurred when the treatment was started at an early stage.

Samuelsson[78] (Sweden) used ECT to destroy colon adenocarcinoma tumors in rats by applying 10–40 mA currents for 1 h and using one to three treatments. Torpedo shaped platinum electrodes were used, with an anode inserted in the centre of the tumor and a cathode placed at the tumor surface. The results obtained by ECT were comparable to those obtained by surgery: 50% of the rats treated with either ECT or surgery were devoid of tumors when killed 7 months after inoculations that induced the tumors.

In other study,[79] Samuelsson and coworkers (Sweden) investigated the efficacy of either ECT or radiotherapy or a combination of the two in destroying experimental colon carcinoma tumor in rats. The combined treatment resulted in tumor growth inhibition, and, total regression was obtained in 75% of the cases; with radiotherapy alone, 75% of the tumors remained. As mentioned above[78] with ECT alone, the regression rate was 50%, the same as surgery alone.

In the work of Griffin and coworkers (U.K.) mentioned[20] in Section I, Introduction and Origins, ECT was used in a *pre-clinical work* on the *normal* livers of male adult outbread white (OBW) rats: the objective here was to observe the patterns of necrosis induced preclinically in the liver itself since it will be the primary normal tissue at risk in the proposed ECT therapy. ECT induced two distinct regions of necrosis, distinguishable histologically and by magnetic resonance imaging: (i) a cylindrical region of primary necrosis centred on the electrode, its volume directly proportional to the charge passed, but greater at the anode than cathode; (ii) a wedge-shaped infarct, apex at the electrode and base extending to the liver edge. The extent of this infarct was greater at the anode than the cathode but showed a sigmoidal relationship to the charge passed during ECT. Results indicated pH changes at the electrodes as likely mediators of tissue injury.

In the work of Chou et al.[22] (U.S.A.) on rats (and mice), ECT, when used in multi-treatments, showed necrosis and regression in all treated tumors; the best 6 months rat tumor cure rate was 75% whereas the best 3 months tumor cure rate in mice was 59.3%.

The tumors under study in this work[22] were radiation induced fibrosarcoma (RIF-1) mouse tumor, and, rat fibrosarcoma.

Some other work from the laboratory of Chou and coworkers[80] (U.S.A.) was on the ECT of breast cancer in rats. Rat breast cancers were initiated by injecting 1×10^6 MTF-7 cells into the right mammary gland fat pad of Fisher 344 female rats. The extent of tumor regression was related to the coulombic dose use in ECT: tumor control rate is less than 40% in the 40C (coulomb) and 60C groups, and, more than 70% in the 80C and 100C groups.

Turler and coworkers[81] (Germany) examined the effectiveness of ECT, in colorectal metastases established in BDIX rats by injection of colon cancer cells under the liver capsule. The tumors thus induced in the liver were subjected to low current level ECT applied via five platinum electrodes (Figure 10). Four different applications were used:

60C cm^{-3}, anode at the centre;
60C cm^{-3}, cathode at the centre;
80C cm^{-3}, anode at the centre;
80C cm^{-3}, cathode at the centre;

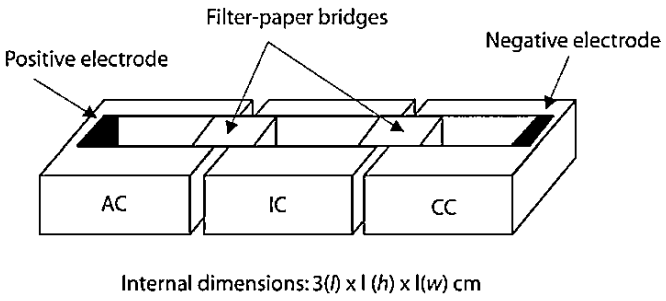


Figure 10. Electrode arrangement during the electrochemical treatment (ECT) with low-level direct current: one platinum electrode is placed at the center of the tumor, four electrodes with opposite polarity are inserted at the border between the tumor and the surrounding liver tissue. Taken from Turler and coworkers.[81]

In the control group, five electrodes were placed without applying any direct current.

Histological examination[81] of the removed livers showed significant destruction of the metastases with localized necroses in all treatment groups: 37% had a complete response rate and 63% a partial response rate. The best results were obtained in the group with an anode at the centre and a current dose of 80 C cm^{-3}. There were no significant necrosis in the control group, as expected.

Colorectal liver tumor deposits in a rat model have also been studied by Maddern and coworkers[82] in Australia. Significant tumor ablation was achieved by ECT conducted by employing several electrode configurations. Tumor necrosis was more complete with electrodes placed on either side of the tumor than with both electrodes placed in the centre of the tumor. In another study[83] on the ECT of liver tumors in rats reported by the same group, i.e., Maddern et al. in Australia, a significant correlation between the current dose and the volume of necrosis, for a given electrode separation, was observed. These workers have also published detailed histological studies on the ECT of normal rat liver and the successful healing of the tissue following the treatment: this work thus established that electrolysis is a safe method for creating defined areas of liver necrosis that heal well with no associated mortality. Such pre-clinical studies are crucial for proving the suitability of ECT for clinical investigations in the Westerns countries.

And, finally, a remarkable study from the same group is on the feasibility of electrolytic ablation of tissue on a rat pancreas;[85] it is worth mentioning that currently, there are no ablative techniques to treat pancreatic cancer in clinical use and this is an aggressive disease with less than 20% of patients suitable for a *curative* surgical resection.

Ten healthy rats were studies in total. Eight rats were treated with variable doses of coulombs and two rats were used as controls. ECT in the rat pancreas produced local necrosis and appears safe and reproducible in the healthy rat tissue; the next step is to examine its feasibility in a pig model pancreas, before evaluating it clinically for pancreatic tumors.

4. Pigs, Rabbits and Dogs

Earlier work on the electrolytic destruction of tissue in pigs is due to Samuelsson and Jönsson[86] in Sweden. The work on pigs has been done on healthy issues (not on cancer tumors) with an aim to establish the nature of tissue destruction, and subsequent healing, caused by ECT. This kind of work on large animals such as pigs, rabbits and dogs etc. is regarded as essential in the occidental countries as a prelude to possible clinical trials on human beings since the parametric confines, safety, toxicity, necrosis and healing patterns etc. can be established on a relatively large mammal.

In the work of Samuelsson and Jönsson,[86] thirty lungs from 26 healthy young pigs, weighing 20 to 30 kg, were subjected to ECT. In most experiments, 8–10 V and 50 to 70 mA were used and platinum electrodes were employed. Both anodic and cathodic sites showed coagulation necrosis and thrombosis in the vessels when examined microscopically. The lesions healed with fibrotic scarring.

In their previous work on the electrolytic destruction of lung tissue in rabbits, the same Swedish group[87] carried out experiments on 130 rabbits weighing between 1.5 and 3 kg. Electrolysis resulted in destruction of tissue both at the anode and the cathode. The anodic lesion has a pH around 3 whereas the pH was about 10 at the cathode. Thrombi could, again, be seen in the vessels within the lesions.

A systematic series of studies on the ECT of pig liver have appeared from the Australian group of Maddern et al. Electrolysis was investigated for generating areas of hepatic necrosis in the pig liver.[88] The lesions healed with time and were associated with minimum morbidity. Further work on 21 pigs showed that during ECT of the liver tissue the electrolytic dose (in coulombs cm^{-3}) correlated with the volume of liver necrosis.[89] This group also established that in addition to the coulombic dose, pH could be used as a realtime monitor to predict more accurately the extent of necrosis.

The normal range (i.e., before ECT) of pH values was 6.5 – 8.7. A pH of less than 6 (at the anode) or more than 9 (at the cathode) reflected total cellular necrosis in those areas.[90] To elucidate the possible complications that may arise during ECT in

the liver region close to a major vein, Maddern and coworkers conducted another study.[91] Under ultrasound control, in six pigs, platinum electrodes were inserted into, or immediately adjacent to, an hepatic vein. As electrolytic dose of 100C was then administered and the evolution of the lesion was monitored using ultrasound. Gas bubbles were seen to enter the hepatic veins or inferior vena cava during treatment in five of the six animals. There were no major complications as a consequence and all animals recovered and remained in a healthy state until they were killed.

This work[88-91] clearly demonstrates that predictable and reproducible necrosis is produced by electrolysis in the pig liver. The treatment causes little or no damage to immediately adjacent liver or major vascular structures.

A very thorough study on the effects of direct current on dog liver has been reported by Li et al.[33] and their main findings have been summarized in the beginning of Section II, Phenomenology of ECT. This work brought out clearly the salient features of ECT such as changes in pH at the anode and the cathode, dehydration of the tissue at the anodic site and oedema at the cathodic site, and, the role of the electrode reactions etc.

It is worth mentioning again that the work reported on mice, hamsters and rats was done on induced tumors and was aimed at destroying them; however, the *in-vivo* investigations published on pigs, rabbits and dogs were carried out in healthy tissues with a view to elucidate the mechanisms, safety, toxicity, parametric limits and techniques etc. of ECT, before applying them to human beings.

V. ELECTROCHEMICAL TREATMENT (ECT) OF TUMORS IN HUMAN BEINGS: CLINICAL STUDIES ON PATIENTS

The earlier clinical studies of ECT in the nineteenth century[1-6] clearly showed the considerable potential of this technique, although standarized protocols could not be established; ECT thus faded away without being integrated into the practice of surgery.

The *rediscovery* of this technique in clinical practice can be traced to the work of Nordenström[11] in Sweden, about a quarter of

a century ago. The adoption of ECT as a common surgical tool to eliminate solid tumors, both cancerous and benign, has occurred in China which, again, can be attributed to the work and inspiration of Nordenström who helped train the initial cadre of Chinese surgeons in the clinical applications of ECT. The leader of this immense Chinese effort is Yu-Ling XIN[29-33] of the China-Japan Friendship Hospital in Beijing, who was previously associated wit the Department of Biomedical Engineering, Hauzhong University of Science and Technology, Hauzhong, China.

1. The New Era in ECT

During the 1950's, Dr. Björn E.W. Nordenström, a Swedish radiologist and surgeon, became curious about the appearance of streaks, spikes and coronas that he saw in the X-ray images of lung tumors: many of his colleagues either saw nothing or attributed these phenomena to artefacts in the images. In 1965 Nordenström began a scientific exploration of these effects. After much experimentation and analysis, he showed that the necrotic centre of the tumor produces a voltage difference between the cancer tumor and the surrounding tissue; this provides the driving force for the electroosmotic movement of water, changes in concentration profiles of ionic components of the tissue, and, the electrophoretic transport of various organic molecules, thus producing detectable structural changes such as coronas, streaks and spikes etc. in the tissue: this is the basis of the so-called Biologically Closed Electric Circuit pathways (BCEC systems) theory of Nordenström elaborated in his book[10] in 1983, and mentioned earlier in this Chapter. The central point of this theory is the presence of electric field profile within the tissue (arising from the voltage difference between the tumor and the surrounding tissue) and its electrochemical consequences viz. electroosmosis, electrophoresis, pH changes etc. The co-transport of charged species (ions and electrons) caused by the presence of electric field creates continuous electric currents within the human body which are moderated by the living system and participate in maintaining equilibrium and healing. He further argued that by augmenting the healing process associated with the naturally occurring BCEC systems in the human body, *electrotherapeutic*

techniques could be developed to treat a variety of diseases including cancer. This led him, in the late 1970s, to treat primary lung cancers by applying current between two platinun wire electrodes. The anode was inserted in the centre of the lung tumor and the cathode was approximately twice the tumor diameter away from the anode; the applied voltage was 10 V (DC). Preliminary work on five patients was reported[11] in 1978, and, in his book published in 1983 he summarized the results on 26 lung tumors in 20 patients.[10] Many of these patients were, for various reasons (e.g., too old, too weak or too sick with other ailments etc), not suitable for conventional treatments such as surgery, radiation or chemotherapy. Regression was observed in 12 out of 26 tumors and no signs of regrowth were detected after a 2–5 year follow-up period.

Further work[14-15] from this group examined the combined effect of ECT and the chemotherapeutic agent, adriamycin. By injecting the electropositively charged adriamycin via the anode in the centre of the tumor, it could be dispersed, by electrophoresis, uniformly into the surrounding tissue. Alternatively, adiramycin could be applied intravenously but its concentration could be augmented within the tumor by placing a cathode within the centre of the tumor. These techniques were applied with beneficial results in a preliminary study of 14 patients with large lung tumors, otherwise incurable.

Nordenström and his coworkers have also applied ECT to breast cancer[17] of a patient. Two years after the treatment, mammograms, taken at six-month intervals, could not detect any remaining tumors.

Other Swedish work on the clinical applications of ECT is due to Samuelsson and coworkers.[92] By applying 80 mA for 2–4 hours to five lung tumors in four patients, they achieved a tumor reduction of 60–80%. Two to three platinum electrodes were used in these experiments, all placed in the tumor.

2. Extensive Clinical Studies in China

In China, ETC has been tried on thousands of patients and the results were summarized by Xin in 1994,[30] 1998,[31] and 2002.[32] Most of this work has been published in Chinese Journals although a special issue of Eur. J. Surgery made this work

available in English in 1994. These investigations do not follow the occidental tradition of randomised and controlled clinical trials and the Western style characterization and follow-up rate of patients; also, the classification of the tumors is not always presented in details. However, these observational studies[30-32] make a compelling case that ECT provides a safe, effective and inexpensive way of treating solid tumors.

Approximately 11000 cancer patients with various kinds of tumors have been treated with electrochemical therapy (ECT). This immense task was accomplished by training 2300 medical doctors through more than 150 ECT training courses between 1987–2001. There have been more than 1300 hospital and clinics in China offering ECT treatment over the past fifteen years. Although most of the treated patients were Chinese, a number of them also came from Europe, Australia, U.S.A., and Japan.

Clinical experience in China[32] indicates that tumor killing area around an electrode is circular with a radius of 1 cm, so that the ideal distance between the anode and the cathode should be approximately 2 cm. The output voltage from DC power source was in the 6 V to 8 V range and the magnitude of current was approximately 60 mA to 80 mA. The electric quantity (total charge or *dosage*) for solid cancer tumors should be 80C to 100C per 1 cm *diameter* (Note that some animal studies[73, 81] propose the use of similar coulombs per cm^3 of the tumor *volume*).

Based on results on thousands of patients who underwent ECT for a variety of tumors, Xin[32] has published extensive data in tabular from. Table 1 shows short-term (6 months) efficacy of ECT whereas Table 2 give the long-term cure rate of various cancers treated with ECT. The results, it is claimed,[32] are equal to or better than those obtained by other techniques, i.e., surgery, radiation and chemotherapy. It is further claimed[32] that ECT can also be applied to patients who cannot be treated by conventional techniques owing to their old age, general poor health, multiple ailments or advances cancers.

The Chinese workers also obtained spectacular results (Table 3) in the treatment of benign tumors of serious nature.[32] In the case of hemangioma (Table 3), a lower electric dosage of 30C to 40C per cm of the tumor is recommended because this benign tumor is very rich in electrolytes and has a very high conductivity; higher coulombic dosages can cause local burns.

Table 1
Short-Term (6 Months) Efficacy of ECT. Objective Remission in 7642 Malignant Tumors Treated with ECT in China,[32] 1987–2001. Responses were Determined by Direct Measurement: CR = Complete Response PR = Partial Response, NC = No Changes, PD = Progressive Disease

Type	Number of cases	CR		PR		NC		PD		CR+PR	
		No.	%	No.	%	No.	%	No.	%	No.	%
Oesophageal cancer	1595	348	21.8	766	48.0	319	20.0	162	10.2	1114	69.8
Lung cancer	1113	412	37.0	445	40.0	155	13.9	101	9.1	857	77.0
Liver cancer	961	240	25.0	427	44.4	210	21.9	84	8.7	667	69.4
Skin cancer	958	611	63.8	347	36.2	0	0	0	0	958	100.0
Breast cancer	644	180	28.0	296	46.0	64	9.9	104	16.1	476	73.9
Cancer of head & face	598	154	25.8	280	46.8	99	16.5	65	10.9	434	72.6
Metastatic lymph node cancer	361	135	37.4	145	40.2	57	15.8	24	6.6	280	77.6
Thyroid cancer	250	89	35.6	110	44.0	36	14.4	15	6.0	199	79.6
Cancer of vulva	237	91	38.2	111	47.0	21	8.8	14	5.9	202	85.2
Melanoma	227	79	34.8	27	11.9	27	11.9	94	41.4	106	46.7
Cancer of chest and abdominal wall	172	44	25.5	81	47.1	29	16.9	18	10.5	125	72.6
Cancer of oral cavity	138	46	33.3	75	54.3	11	8.0	6	4.4	121	87.6
Rhabdomyosarcoma	133	27	20.3	49	36.8	38	28.6	19	14.3	76	57.1
Parotid cancer	84	28	33.3	46	54.8	6	7.1	4	4.8	74	88.1
Throat cancer	21	9	42.9	9	42.9	2	9.5	1	4.7	18	85.7
Prostate cancer	20	8	40.0	7	35.0	3	15.0	2	10.0	15	75.0
Others	130	39	30.0	51	39.2	20	15.4	2	15.4	90	69.2
Sum total	**7642**	**2540**	**33.2**	**3272**	**42.8**	**1097**	**14.4**	**733**	**9.6**	**5812**	**76.0**

Table 2
Long-Term Effectiveness of ECT Determined as the 1–5-Years Survival Rate in 7642 Cases of Malignant Tumors Treated in China,[32] 1987–2001

Type	Number of cases	1 year		2 years		3 years		4 years		5 years	
		No.	%	No.	%	No.	%	No.	%	No.	%
Oesophageal cancer	1595	1285	80.6	969	60.8	483	30.3	233	14.6	205	12.9
Lung cancer	1113	1063	95.5	933	83.8	746	67.0	600	53.9	432	38.8
Liver cancer	961	771	80.2	577	60.0	209	21.7	184	19.1	145	15.1
Skin cancer	958	958	100.0	958	100.0	890	92.9	890	92.9	767	80.1
Breast cancer	644	618	96.0	568	88.2	404	62.7	404	62.7	323	50.2
Cancer of head & face	598	501	83.8	428	71.6	338	56.5	327	54.7	245	41.0
Metastatic lymph node cancer	361	336	93.1	295	81.7	258	71.5	184	51.0	139	38.5
Thyroid cancer	250	242	96.8	232	92.8	192	76.8	191	76.4	133	53.2
Cancer of vulva	237	223	94.1	209	88.2	188	79.3	188	79.3	104	43.9
Melanoma	227	194	85.5	157	69.2	91	40.1	19	8.4	0	0
Cancer of chest and abdominal wall	172	140	81.4	130	75.6	115	66.9	102	59.3	74	43.0
Cancer of oral cavity	138	129	93.5	124	89.9	111	80.4	99	71.7	85	61.6
Rhabdomyosarcoma	133	121	91.0	88	66.2	51	38.3	9	6.8	0	0
Parotid cancer	84	78	92.9	76	90.5	67	79.8	59	70.2	45	53.6
Throat cancer	21	21	100.0	21	100.0	19	90.4	17	81.0	14	66.7
Prostate cancer	20	20	100.0	19	95.0	18	90.0	16	80.0	10	50.0
Others	130	124	95.4	99	76.2	98	75.4	71	54.6	31	23.8
Sum Total	**7642**	**6824**	**89.3**	**5883**	**77.0**	**4278**	**56.0**	**3593**	**47.0**	**2752**	**36.0**

Table 3
Long-Term Survival Rates of 1369 Cases of Benign Tumors Treated by ECT in China,[32] 1987–2001

Type	Number of cases	1 year		2 years		3 years	
		No.	%	No.	%	No.	%
Hemangioma[a]	874	874	100.0	874	100.0	874	100.0
Thyroid tumor	116	116	100.0	116	100.0	116	100.0
Thyroid cyst	102	102	100.0	102	100.0	102	100.0
Prostate hypertrophy[b]	191	191	100.0	189	98.9	156	81.7
Keloid[c]	86	86	100.0	86	100.0	86	100.0
Sum total	**1369**	**1369**	**100.0**	**1367**	**99.8**	**1334**	**97.4**

[a] Hemangioma is a purplish or reddish *benign* tumor consisting of a network of interconnecting capillaries.
[b] Prostate hypertrophy is benign enlargement of the prostate.
[c] Keloid is a thick scar resulting from excessive growth of fibrous tissue and behaves as a benign tumor.

For lung cancer, some data comparing ECT with other conventional techniques have been reported[29] and are summarized in Table 4.

Table 4
Long-Term (5 Years) Survival Rates of Lung Cancer: Comparison of ECT with other Conventional Techniques[a]

Technique	% Survival at 5 years
Surgery	42.5
Chemotherapy	16.5
Radiation and chemotherapy	11.0
ECT	25.5

[a] The above data are based on the review of studies reported by Xin et al.[29] Note that ECT was conducted on patients who were either in advanced stages of lung cancer and/or had failed radiotherapy and/or chemotherapy.

3. Clinical Studies in Other Countries

As stated earlier, the clinical investigations in other countries are few and far between since surgical practice treats ECT as *experimental* therapy because protocols based on randomised studies are not established yet.

The Japanese work has been carried out by the group of Matsushima[28,93] in the Department of Surgery, Tokyo Medical College. They found that a combination of ECT with chemotherapy gave the best results for treating patients with lung cancers; especially, the anti-cancer drugs tend to accumulate around the electrodes when direct current is passed. In a clinical study on 26 patients who received ECT either with or without the anti-cancer drugs, a decrease in tumor size was observed in more than 70% the cases. In two cases, the use of ECT alone resulted in complete regression of the tumors.[28]

Clinical work from Slovenia[94] involves treatment of five patients with melanoma skin tumors; ECT was combined with systemic chemo- and/or immunotherapy. Tumor mass reductions were observed but no complete responses.

A very large study jointly conducted by investigators in France, U.S.A. and Slovenia, and, involving five cancer centres has explored the effectiveness of electric current and chemotherapy combination in treating cutaneous and subcutaneous malignant tumors.[95]

The current was applied in the form of electrical pulses in order to permeabilize the membranes of the tumor cells for the entry of the chemotherapeutic agent *bleomycin* which is a very potent cytotoxic molecule. Clinical complete responses were achieved in 56.4% of the tumors and partial responses were observed in 28.9% of the tumors. This work is thus not strictly electrochemical treatment in the sense of Nordenström[18] and Xin[32] but is rather chemotherapy aided by electrochemical-driven movement of ions, molecules and drugs etc. (e.g., by electroosmosis, electrophoresis) into the tissue regions targeted for necrosis, as in several studies[28,51,93,94] described earlier within this chapter.

The American partners in this large effort[95] have reported, separately, on the novel electrode designs for administering this electrochemotherapy: electrical pulses permeabilize, or

electroporate, tumor cell membranes which allows the chemotherapeutic agent greater access to intracellular site of action and consistently provides improved responses relative to treatment with the drug alone.[96] Other similar work from the U.S.A. may also be mentioned.[97]

In a joint study[98] from the U.K. and Australia, metastatic colorectal cancer lesions in the liver of a 58 year old man were electrolytically ablated. Follow-up CT scan suggested complete tumor regression. The Australian team of Maddern and coworkers[99] performed ECT combined with surgery on nine patients with hepatic deposits from colorectal carcinoma; within the electrolytically treated area, seven patients had no radiological evidence of recurrence at a median follow-up of 9 (range 6–43) months whereas local recurrence was detected in two patients. Six of the nine patients had metastases elsewhere in the liver with four having extrahepatic metastases. Three patients remain tumor free. Three patients died. The median survival was 17 (range 9–24) months from the time of treatment. A review article comparing electrolysis with other local ablative treatments for non-resectable colorectal liver metastases has been jointly published by the U.K./Australian team.[100] It was concluded that electrolysis offers potential benefits over other techniques (e.g., cryotherapy, microwave coagulation therapy, radiofrequency ablation and lase-induced thermotherapy) that include the apparent ability to safely and effectively treat lesions abutting major hepatic structures, and, lack of a systemic inflammatory reaction following electrolytic ablation.

VI. CONCLUSIONS

The salient points in the electrochemical treatment (ECT) of tumors, based on the detailed descriptions in this chapter, may be summarized as follows:

1. The approach involves a passage of direct electric current through the tumor, thus destroying the diseased tissue through a number of electrochemical events.
2. As regards the mechanism of ECT, the passage of current through the tumor gives hydrogen evolution at the cathode causing cavitation (and hypoxia) of the tissue and chlorine

evolution at the anode that results in bleaching and destruction of the tissue in the anode area. Further contributions to the tissue necrosis are made by; (i) the changes in pH caused by the electrode reactions; (ii) electroosmotic movement of water from the anode to the cathode causing dryness of the tissue in the anodic area and its oedema near the cathode area. All these events profoundly perturb the metabolism of the cancerous tissue and destroy its structure and function, thus leading to the annihilation of the tumor. The passage of electric current also leads to micro-thromboses in the capillaries feeding the tumor, reinforcing further the destruction of its tissue.

3. The methodology used is quite simple and empirical in nature: an application of 6–10 volts (DC) between two electrodes about 1–2 cm apart within the tissue, causing a passage of 60–100 mA (with wire electrode area about 0.2 cm^2) and the experiment lasts from one to two hours. In general, an array of anodes and cathodes is used to cover the entire solid tumor and the bordering tissue. A more rigorous and quantitative methodology is needed in order to increase the efficacy and accuracy of ECT.
4. The ECT has been applied with much success both in animals and human beings. It has been demonstrated to be a safe, effective and inexpensive procedure.

For it to be applied more extensively at the clinical level, one needs to develop standardized commercial instrumentation and its utility has to be demonstrated in more controlled and rigorous pre-clinical and clinical studies in terms of the accepted surgical norms in North America and Western Europe.

REFERENCES

[1]J. Althaus, *British Med. J.* **Nov. 15** (1875) 606.
[2]F. H. Martin, *J. Amer. Med. Ass.* (*JAMA*) **8** (1886) 61; idem, ibid, **8** (1866) 85.
[3]J. N. Freeman, *J. Amer. Med. Ass*. (*JAMA*) **8** (1886) 90.
[4]Editorial, *British Med. J.* **Dec. 6** (1873) 667.
[5]E. Van de Warker, *J. Amer. Med. Ass*. (*JAMA*) **8** (1886) 88.
[6]R. Newman, *J. Amer. Med. Ass.* (*JAMA*) **8** (1886) 88.
[7]B. W. Watson, *Med. Sci. Res*. **19** (1991) 103.
[8]N. Melczer and J. Kiss, *Nature* **4571** (1957) 1177.

[9]C. E. Humphrey and E. H. Seal, *Science* **130** (1959) 388.
[10]B. E. W. Nordenström, *Biologically-closed electric circuits*, Nordic Medical Publication, Stockholm, 1983.
[11]B. E. W. Nordenström, *IRCS Medical Science* **6** (1978) 537.
[12]B. E. W. Nordenström, *Ann. Radiol*. **28** (1985) 128.
[13]B. E. W. Nordenström, *Amer. J. Clin. Oncol* (*CCT*) **12**(6) (1989) 530.
[14]B. E. W. Nordenström, Amer. J. Clin. Oncol (*CCT*) **13**(1) (1990) 75.
[15]S. Eksborg, B. E. W. Nordenström, and H. Beving, *Amer. J. Clin. Oncol.* (*CCT*) **13**(2) (1990) 164.
[16]H. Beving, S. Eksborg and B. E. W. Nordenström, *Amer. J. Clin. Oncol* (*CCT*) **13**(2) (1990) 167.
[17]E. Azavedo, G. Svane and B. E. W. Nordenström, *Clinical Radiology* **43** (1991) 84.
[18]B. E. W. Nordenström, *Eur. J. Surg, Suppl*. **574** (1994) 93.
[19]E. Nilsson, H. V. Eulor, J. Berendson, A. Thorne, P. Wersäll, I. Näslund, A-S. Lagersted, K. Narfström, and J. M. Olsson, *Biolectrochemistry* **1** (2001) 1.
[20]D. T. Griffin, N. F. J. Dodd, S. Zhao, B. R. Pullan, and J. V. Moore, *British J. Cancer* **72** (1995) 31.
[21]C. K. Chou, *Bioelectromagnetics* **18** (1997) 1.
[22]C. K. Chou, J. A. Mcdougall, C. Ahn, and N. Vora, *Bioelectromagnetics* **18**, (1997) 14.
[23]B. G. Fosh, J. G. Finch, M. Lea, C. Black, S. Wong, S. A. Wemyss-Holden, and G. J. Maddern, *British J. Surg*. **89** (2002) 999.
[24]P. S. Baxter, S. A. Wemyss-Holden, A. R. Dennison, and G. J. Maddern, *Aust. NZ. J. Surg*. **68** (1998) 637.
[25]A. K. Vijh, *J. Mater Sci: Materials in Medicine* **10** (1999) 419; A. K. Vijh, *Drying Tech*. **17** (1999) 585.
[26]S. Srinivasan, G. L. Cahan and G. E. Stoner, in *Electrochemistry: The Past Thirty Years and the Next Thirty Years*, Chapter 4, Ed. by H. Bloom and F. Gutman, Plenum Press, New York, 1977.
[27]*Eur. J. Surgery, Suppl***. 574** (1994), the entire issue.
[28]Y. Matsushima, E. Takahashi, K. Hagiwara, C. Konaka, H. Miura, H. Kato, and Y. Koshiishi, *Eur. J. Surg., Suppl***. 574** (1994) 59.
[29]Y. L. Xin, F. S. Xue, B. S. Ge., F. R. Zhao, B. Shi, and W. Zhang, *Bioelectromagnetics*, **18** (1997) 8.
[30]Y. L. Xin, *Eur. J. Surg., Suppl***. 574** (1994) 25.
[31]Y. L. Xin, *Preprints from the 2nd International Symposium on Electrochemical Treatment of Cancer*, Beijing, 27-30 Sept. 1998, p. 81-92.
[32]Y. L. Xin, *Journal of the IABC in Medicine and Biology*, **1**(1), (2002) 24. (Note: IABC denotes International Association For Biologically-Closed Electric Circuits (BCEC) in Medicine and Biology].
[33]K. H. Li, Y. L. Xin, Y. N. Gu, and B. L. Xu, *J. Bioelectromagnetics*, **18** (1997) 2.
[34]C. K. Chou and R. L. Ren, in *Complementary/Alternative Medicine—An Evidence-Based Approach*, Chapter 2, Ed. by J. W. Spencer and J. J. Jacobs, Mosby Publishers, Boston, 1999.
[35]Y. L. Xin, *Eur. J. Surg., Suppl*. **574** (1994) 31.
[36]J. O'M Bockris (retired from Texas A & M University, College Station, Texas) personal communication to A. K. Vijh (Sept. 3, 2000).

[37]N. L. Vora, *Electrochemical Treatment of Breast Cancer with Direct Current*, U.S. Army Medical Research and Material Command, Fort Detrick, Maryland 21702-5012, U.S., Report on Award Number DAMD17-96-1-6184 (June 1999).
[38]S. L. David, D. R. Absolom, C. R. Smith, J. Gams and M. A. Herbert, *Cancer Research* **45** (1985) 5625.
[39]B. T. Mortensen and J. Bojsen, *J. Biomed. Engg*. **4** (1982) 100.
[40]D. Miklavčič, G. Serša, B. Kryžanowski, S. Novakovič, F. Bobanovič, R. Golauh, and L. Vodovnik, *Bioelectrochem. and Bioenergetics* **30** (1993) 209.
[41]D. Miklavčič, G. Serša, S. Novakovič and S. Reberšek, *J. Bioelectr*. **9** (1990) 133.
[42]M. B. Habal and M. K. Scauble, *Surg. Forum* **18** (1967) 88.
[43]F. Shirazi and B. C. Towe, *Proc. 8th BRAGS*, Washington, D.C., Oct. 9-12, (1988) p. 60.
[44]M. P. Abbracchio, J. D. Heck, and M. Costa, *Carcinogenesis* **3** (1982) 175.
[45]R. K. Jain, *Sci. Amer*. **271**(1) (1994) 58.
[46]S. Ohki, in *Comprehensive Treatise of Electrochemistry*, Vol. 10, Ed. by S. Srinivasan, Y. A. Chizmadzhev, J. O'M Bockris, B. E. Conway, and E. Yeager, Plenum Press, New York,1985, pp. 188, 199.
[47]B. E. Conway, *Theory and Principles of Electrodes Processes*, Ronald Press, New York, 1965.
[48]D. H. Everett, *Basic Principles of Colloid Science*, The Royal Society of Chemistry, London, England, 1988.
[49]A. K. Vijh, *Drying Technology—An International Journal* **13** (1995) 215.
[50]A. K. Vijh, in *Modern Aspects of Electrochemistry*, Ed. by J. O'M Bockris, B. E. Conway, and R. E. White, Plenum Press, New York, 1999, Vol. 32, Chap. 4.
[51]G. Serša, S. Novakovic, and D. Miklavčič, *Cancer Lett*. **69** (1993) 81.
[52]I. Ito and S. Hashimoto, *Granto Kagaku Ryoho* **16** (1989) 1405.
[53]G. Serša, R. Golauh, and D. Miklavčič, *Anti-Cancer Drugs* **5** (1994) 69.
[54]G. Serša and D. Miklavčič, *Mol. Biother* **2** (1990) 165.
[55]G. Serša, D. Miklavčič, *Radiol. Oncol*. **27** (1993) 280.
[56]G. Serša, D. Miklavčič, U. Batista, S. Novakovič, F. Bobanovič, and L. Vodovnik, *Anti-Cancer Drugs* **3** (1992) 253.
[57]E. Nilsson, J. Berendson, and E. Fontes, *J. Electroanal. Chem*. **460** (1999) 88.
[58]E. Nilsson, J. Berendson, and E. Fontes, *J. Appl. Electrochem*. **30** (2000) 1321.
[59]J. G. Finch, B. Fosh, A. Anthony, E. Slimani, M. Texler, D.P. Berry, A.R. Dennison, and G. J. Maddern, *Clinical Science* **102** (2002) 389.
[60]M. Lyte, J. E. Gannon, and G. D. O'Clock, Jr., *J. Natl. Cancer Inst*. **83** (1991) 116.
[61]Y. Yen, J-R. Li, B-S. Zhou, F. Rojas, J. Yu, and C.K. Chou, *Bioelectromagnetics* **20** (1999) 34.
[62]V. F. Veija, C. Holandino, M. L. Rodrigues, M. A. M. Capella, S. Menezes, and C.S. Alviano, *Bioelectomagnetics* **21** (2000) 597.
[63]G. Kannagi, *Glycoconj. J*. **14** (1997) 577.
[64]J. Alper, *Science* **301** (2003) 159.
[65]H. Ito, N. Shiegmatsu, T. Nakayama, I. Nishiguchi, Y. Ando and S. Hashimoto, *J. Jap. Soc. Cancer Therapy* **23** (1988) 696.
[66]D. N. Morris, A. A. Marino, and E. Gonzalez, *J. Surgical Res*. **53** (1992) 306.
[67]D. T. Griffin, N. J. F. Dodd, J. V. Moore, B. R. Pullan, and J. V. Taylor, *British J. Cancer* **69** (1994) 875.

[68]D. Miklavčič, G. Serša, S. Novakovič, and S. Reberšek, *J. Bioelectricity* **9** (1990) 133.
[69]D. Miklavčič, L. Vodovnik, F. Bobanovič, S. Reberšek, G. Serša, S. Novakovič, and R. Golouh, *Electro-Magnetobiol*. **11** (1992) 109.
[70]D. Miklavčič, G. Serša, M. Kryžanowski, S. Novakovič, F. Bobanovič, R. Golouh, and L. Vodovnik, *Bioelectrochem Bioenerg*. **30** (1993) 209.
[71]G. Serša and D. Miklavčič, *Reg. Cancer Treat*. **1** (1993) 31.
[72]D. Miklavčič, A. Fajgelj, and G. Serša, *Bioelectrochem. Bioenerg*. **35** (1994) 93.
[73]L. B. Cabrales, H. C. Ciria, R. B. Bruzon, M. S. Quevedo, R. H. Aldana, L. B. de Oca, M. F. Salas, and O. de la Guardia Peña, *Bioelectromagnetics* **22** (2001) 316.
[74]M. K. Schauble, M. B. Habal, and H. D. Gullick, *Arch. Pathol. Lab. Med.* **101** (1977) 294.
[75]S. L. David, D. R. Absolom, C. R. Smith, J. Gams, and M. A. Herbert, *Cancer Res*. **45** (1985) 5625.
[76]B. Gardner, A. Sterling, and J. Brown, *Surgery* **62** (1967) 361.
[77]M. . Habal, J. Biomed. *Mater. Res*. **14** (1980) 789.
[78]L. Samuelsson, *Acta Radio*. **22** (1981) 129.
[79]L. Samuelsson, L. Jönsson, I-L. Lamm, C. J. Lindén, and S.B. Ewers, *Acta Radiol.* **32** (1990) 1978.
[80]R. L. Ren, N. Vora, F. Yang, J. Longmate, W. Wang, H. Sun , J-R. Li, L. Weiss, C. Staud, J. A. McDougall, and C. K. Chou, *Bioelectromagnetics* **22** (2001) 205.
[81]A. Turler, H. Shaefer, N. Schaefer, M. Wagner, D. Maintz, J.C. Qiao and A.H. Hoelscher, *Bioelectromagnetics* **21** (2000) 395.
[82]S. A Wemyss-Holden, G. S. M. Robertson, P. de la M. Hall, A. R. Dennison, and G. J. Maddern, *Digestive Diseases* **18** (2000) 50.
[83]G. S. M. Robertson, S. A. Wemyss-Holden, A. R. Dennison, P. de la M. Hall, P. Baxter, and G. N. Maddern, *British J. Surgery* **85** (1998) 1212.
[84]S. A. Wemyss-Holden, G. S. M. Robertson, A. R. Dennison, P. de la M. Hall, J. C. Fothergill, B. Jones, and G. J. Maddern, *J. Surg. Res.* **93** (2000) 55.
[85]B. G. Fosh, J. G. Finch, A. A. Anthony, M. Texler, and G. J. Maddern, BMC *Gastroenterology* **1** (2001) 9.
[86]L. Samuelsson and L. Jönsson, *Acta Radiologica Diagnosis* **22** (1981) 9.
[87]L. Samuelsson, T. Olin and N. O. Berg, *Acta Radiologica Diagnosis* **21** (1980) 447.
[88]S. A. Wemyss-Holden, G. S. M. Robertson, A. R. Dennison, P. S. Vanderzon, P. de la M. Hall, and G. J. Maddern, *Clinical Science* **98** (2000) 561.
[89]S. A. Wemyss-Holden, A. R. Dennison, G. J. Finch, P. de la M. Hall, and G. J. Maddern, *British J. Surgery* **89** (2002) 579.
[90]J. G. Finch, B. Fosh, A. Anthony, E. Slimani, M. Texler, D. P. Berry, A. R. Dennison, and G. J. Maddern, *Clinical Science* **102** (2002) 389.
[91]S. A. Wemyss-Holden, P. de la M. Hall, G. S. M. Robertson, A. R. Dennison, P. A. Vanderzon, and G. J. Maddern, *ANZ J. Surg.* **70** (2000) 607.
[92]L. Samuelsson, L. Jönsson, and E. Ståhl, *Radiologe* **23** (1983) 284.
[93]Y. Matsushima, R. Amemiya, J-S. Liu, E. Tajika, H. Takakura, K. Oho, and Y. Hayata*, J. Jap. Soc. Cancer Ther*. **24** (1989) 2341.
[94]A. Plesničca, G. Serša, L. Vodovnik, J. Jancar, L. Zaletel-Kragelj, and S. Plesničar, *Eur. J. Surg. Suppl*. **554** (1994) 45.
[95]L. M. Mir, L. F. Glass, G. Serša, J. Teissié, C. Domenge, D. Miklavčič, J. J. Jaroszeski, S. Orlowski, D. S. Reintgen, Z. Rudolf, M. Belehradek, R. Gilbert,

M-P. Rols, J. Belehradek, J. M. Bachand, R. DeConti, B. Štabuc, M. Čemažar, P. Coninx, and R. Heller, *British J. Surgery* **77** (1998) 2336.
[96]R. A. Gilbert, N. J. Jaroszeski, and R. Heller, *Biochmica and Biophysica Acta* **1334** (1997) 9.
[97]S. B. Dev and G. A. Hofmann, *Cancer Treatment Reviews* **20** (1994) 105.
[98]D. P. Berry, A. R. Dennison, R. Ward, and G. J. Maddern, *Dig. Surgery* **17** (2000) 518.
[99]B. G. Fosh, J. G. Finch, M. Lea, C. Black, S. Wong, S. A. Wemyss-Holden, and G. J. Maddern, *Brisith J. Surgery* **89** (2002) 999.
[100]B. D. Teague, S. A. Wemyss-Holden, B. G. Fosh, A. R. Dennison, and G. J. Maddern, *ANZ J. Surg.* **72** (2002) 137.

Index

AC impedance spectroscopy, 190
Active islands, inactive electrodes, 213
Adsorption isotherms, 183
Analysis, small-signal, 162
Anodic deposit formation, 145
Anodic processes, formation of PS, 104
Aqueous polarization, 25
Array of anodes and cathodes, 270
Atomic force microscopy (AFM), 175
Bockris, 42
Born equation, 46
Brownian movement, 39
Capillary condensation, 183
Cell metabolism, 235
Chandrasekhar, 43
Clinical applications, ECT, 263
Continuous-flow stirred tank electrochemical reactor, 152
Conway, 42
Copper dissolution, anodic, 146
Corrosion, fractal approach, 169
Cottrell behavior, 214
Cottrell equation, 185, 193
Coulostatic analysis, 162
Crystalline silicon, anodic dissolution, 65
Current burst theory, dissolution, 69
Current density, porous silicon, 76
Current transient, fractality, 197
Cyclic voltammetry, 189
DC treatment, target tumor cells, 251
Deposit dissolution process, 145
Diffuse double layer, 244
Diffusion coefficients, 14
Dipole moments, 25
Discrete molecular models, 17, 22
Dissolution rate, mass transfer, 68
Dissolution valence, 100
Dissolution valence, effective, 75
Dissolution, chemical, 75
Doped n-Si, 79
Doped p-Si, 79
ECT method, standarized treatment, 237
Electric field, human body, 262

ionic concentration, 245
tumor, 262
Electrical current,
tumor treatment, 231
Electrical gradient,
water flow, 247
Electrochemical double layer,
cell interstitium interface, 243
charge neutrality, 245
Electrochemical engineering,
Markov theory, 136
Electrochemical processes,
non-diffusion-limited, 220
Electrochemical reactions,
dissolution of silicon, 103
Electrochemical techniques,
diffusion-limited, 169
Electrochemical treatment (ECT),
of tumors, 232, 268
low-level direct current, 258
Electrochemical treatment,
salient points, 269
Electrode surface processes, 142
Electrodeposition,
fractal approach, 169
Electrolytic cell,
flow circulation, 160
Electrolyzer, 148
Electrolyzer,
parasitic reactions, 159
Electroosmosis,
interstitial channels, 240
Electroosmosis,
tissue, 240
Electroosmotic dewatering,
zeta potential, 249
Electroosmtic Dewatering (EOD), 241
Electrophoretic transport,
organic molecules, 262
Electropolishing, 73
Electrostatic solvation free energy,
qunatum-mechanical, 30
Erlang formula, 140, 143
Failed cells,
Markovian analysis, 153
First solvation shell, 2
Fowler, 42
Fractal dimension,
surface, 169
Fractal geometry,
Mandelbrot, 170
Fractals, self-similar, 171
nonrandom, 170
random, 170
self-affine, 169, 170
Free energies of solvation in, 11, 24
Free energies of solvation in water, 23
Free energies, cavitation, van der Waals, 33
Free energy perturbation procedure (FEP), 20
Free energy,
hydration, 33
Fuel-cell electrocatalyst, 214
Fullerene,
solubility, 14
Gas adsorption, molecular probe method, 181
General interaction properties function (GIPF), 8
Gibbs free energy, 20
Gibbs-Helmholtz equation, 46
Growth rate, 76
Hartree-Fock calculations, 6
Hellmann-Feynman theorem, 1

Helmholtz layer, 100
Helmholtz-Smoluchowski relation, 245
HF solution, 120
Highest occupied molecular orbital (HOMO), 54
Hildebrand, solubility parameter, 51
Hirschfelder, 1
Hydration, ionic enthalpies, 45
free energies, 45
Hydrogen bonding, 10
evolution, 71, 235
evolution, cavitation, 269
Image analysis methods, 174
Interaction energies, 1
Interfaces, electrochemistry, fractal approach, 167
Intramolecular interactions, 18
Ionic solvation, continuum solvent models, 44
discrete molecular solvent models, 43
Kolmogorov, 139
Langevin dipoles, 39
Laplace transforms, 187
Layer thickness, 143
Lennard-Jones potential, 17
Linear Sweep/Cyclic Voltammetry, 188
Lithium transport, cell-impedance-controlled, 220
Lowest unoccupied molecular orbital (LUMO), 54
Mandelbrot, fractal geometry, 167
Markov matrix, 138
Markov process, 136, 139
Markov chains, processes, 135
Mass fractal, 169
Mathematical modelling, ECT, 250
Mechanistic aspects, electrochemical treatment, 240
Mesoporous carbons, 219
Metallic electrodes, tumor destruction, 239
Models of solvent, hybrid/intermediate, 38
Molecular dynamics, 2
electrostatic potential, 3
volume, 53
Monte Carlo simulations, of solutions, 2
Multi-array electrode, ECT effect, 255
Nernst equation, redox couples, 188
Nitroaromatics, 13
Nonpolar solvents, 11
Ohmic resistance, substrate, 107
Onsager, 28
Organic solutions, anhydrous, 103
Oxide film thickness, 114
Para-dinitrobenzene, 7
Partition coefficients, 12
Peak-current method, 189
Performance, Markovian modeling, 148
Perimeter-area method, 178
Perturbation theory, 2
Phenomenology of ECT, 246
Photo holes, 119
Physical adsorption, 1
Poisson distribution, 140

Polarization, solute molecule, 31
Pore
  arrays, 85
  branching, 89
  diameter, 82
  diameter, distribution of, 87
  diameter, interpore spacing, 86
  orientation, 87
Porous electrode, 218
Porous silicon, 71
  formation, 71
  morphological characteristics of, 97
  morphology, 65
Positive charge,
  extra-cellular fluid, 241
Potential gradients,
  double layer, 243
Potential,
  effect of current and HF Concentration, 83
Potentiostatic current transient method, 185
Probability distribution, 157
Pt/polished Al_2O_3 electrode, 201
Randles-Sevcik equation, 188
Random walk, Ion, 142
Reaction field, 27
Reaction kinetics,
  anodic, 98
Rough surfaces,
  fractal geometry, 167
Roughness parameter, 174
Scanning electron microscopy (SEM),
  transmission electron microscopy, 174
Scanninng tunneling microscopy (STM), 174
Self-Affine fractal, 172
Silicon, 91
Silicon oxide,
  dissolution, 112
Silicon substrate, 74
Silicon/Electrolyte interface, 99
Solid tumor, 270
Solute Hammiltonian, 25
Solute-solvent interactions, 1, 9
Solvent effects, 15
Solvent,
  continuum models, 26
Solvent-solvent interactions, 51
Space charge layer,
  semiconductor, 106
Sputter-deposited Pt films, 204
Sputter-deposited, Pt film electrodes, 201
Structure-activity relationships, 12
Supercritical solvents,
  solubility in, 15
Surface fractal dimension,
  interfaces, 174
Surface lattice structure,
  silicon, 110
Surface tension, solvent, 33
Switching-Circuit operations,
  analysis of, 157
Symmetry-adapted perturbation theory (SAPT), 2
Tank electrolyzer, 150
TEM image analysis method, 218
Thermal emission,
  of holes, 73

Transition layer, 92
Triangulation method, fractals, 175
Zener tunnelling, 73
Zeta potential, electroosmotic velocity, 245