*Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl*

**Physics and Chemistry of Interfaces**

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

*Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl*

# Physics and Chemistry of Interfaces

WILEY-VCH

WILEY-VCH GmbH & Co. KGaA

**Authors**

Hans-Jürgen Butt
MPI für Polymerforschung Mainz
e-mail: butt@mpip-mainz.mpg.de

Karlheinz Graf
MPI für Polymerforschung Mainz
e-mail: grafk@mpip-mainz.mpg.de

Michael Kappl
MPI für Polymerforschung Mainz
e-mail: kappl@mpip-mainz.mpg.de

**Cover Pictures**
The left picture shows aggregates of silicon oxide particles with a diameter of 0.9 µm (see example 1.1). At the bottom an atomic force microscope image of cylindrical CTAB micelles adsorbed to gold(111) is shown (see example 12.3, width: 200 nm). The right image was also obtained by atomic force microscopy. It shows the surface of a self-assembled monolayer of long-chain alkylthiols on gold(111) (see fig. 10.2, width: 3.2 nm).

This book was carefully produced. Nevertheless, authors and publisher do not warrant the information contained therein to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate

**Library of Congress Card No. applied for**

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

**Bibliographic information published by Die Deutsche Bibliothek**
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

© 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not be considered unprotected by law.

printed in the Federal Republic of Germany
printed on acid-free paper

**Composition** Uwe Krieg, Berlin
**Printing** betz-druck GmbH, Darmstadt
**Bookbinding** Litges&Dopf Buchbinderei GmbH, Heppenheim

**ISBN** 3-527-40413-9

# Preface

Interface science has changed significantly during the last 10–15 years. This is partially due to scientific breakthroughs. For example, the invention of scanning probe microscopy and refined diffraction methods allow us to look at interfaces under "wet" conditions with unprecedented accuracy. This change is also due to the greatly increased community of interfacial scientists. One reason is certainly the increased relevance of micro- and nanotechnology, including lab-on-chip technology, microfluids, and biochips. Objects in the micro- and nanoworld are dominated by surface effects rather than gravitation or inertia. Therefore, surface science is the basis for nanotechnology.

The expansion of the community is correlated with an increased interdisciplinarity. Traditionally the community tended to be split into "dry" surface scientists (mainly physicists working under ultrahigh vacuum conditions) and "wet" surface scientists (mainly colloid chemists). In addition, engineers dealing with applications like coatings, adhesion, or lubrication, formed a third community. This differentiation is significantly less pronounced and interface science has become a really interdisciplinary field of research including, for example, chemical engineering and biology.

This development motivated us to write this textbook. It is a general introduction to surface and interface science. It focuses on basic concepts rather than specific details, on understanding rather than learning facts. The most important techniques and methods are introduced. The book reflects that interfacial science is a diverse field of research. Several classical scientific or engineering disciplines are involved. It contains basic science and applied topics such as wetting, friction, and lubrication. Many textbooks concentrate on certain aspects of surface science such as techniques involving ultrahigh vacuum or classical "wet" colloid chemistry. We tried to include all aspects because we feel that for a good understanding of interfaces, a comprehensive introduction is required.

Our manuscript is based on lectures given at the universities of Siegen and Mainz. It addresses (1) advanced students of engineering, chemistry, physics, biology, and related subjects and (2) scientists in academia or industry, who are not yet specialists in surface science but want to get a solid background knowledge of the subject. The level is introductory for scientists and engineers who have a basic knowledge of the natural sciences and mathematics. Certainly no advanced level of mathematics is required. When looking through the pages of this book you will see a substantial number of equations. Please do not be scared! We preferred to give all transformations explicitly rather than writing "as can easily be seen" and stating the result. Chapter "Thermodynamics of Interfaces" is the only exception. To appreciate it a basic knowledge of thermodynamics is required. You can skip and still be able to follow the rest. In this case please read and try to get an intuitive understanding of what surface excess is (Section 3.1) and what the Gibbs adsorption equation implies (Section 3.4.2).

A number of problems with solutions were included to allow for private studies. If not mentioned otherwise, the temperature was assumed to be 25°C. At the end of each chapter the most important equations, facts, and phenomena are summarized to given students a chance to concentrate on important issues and help instructors preparing exams.

One of the main problems when writing a textbook is to limit its content. We tried hard to keep the volume within the scope of one advanced course of roughly 15 weeks, one day per week. Unfortunately, this means that certain topics had to be cut short or even left out completely. Statistical mechanics, heterogeneous catalysis, and polymers at surfaces are issues which could have been expanded.

This book certainly contains errors. Even after having it read by different people independently, this is unavoidable. If you find an error, please write us a letter (Max-Planck-Institute for Polymer Research, Ackermannweg, 55128 Mainz, Germany) or an e-mail (butt@mpip-mainz.mpg.de) so that we can correct it and do not confuse more students.

We are indebted to several people who helped us collecting information, preparing, and critically reading this manuscript. In particular we would like thank R. von Klitzing, C. Lorenz, C. Stubenrauch, D. Vollmer, J. Wölk, R. Wolff, K. Beneke, J. Blum, M. Böhm, E. Bonaccurso, P. Broekmann, G. Glasser, G. Gompper, M. Grunze, J. Gutmann, L. Heim, M. Hillebrand, T. Jenkins, X. Jiang, U. Jonas, R. Jordan, I. Lieberwirth, G. Liger-Belair, M. Lösche, E. Meyer, P. Müller-Buschbaum, T. Nagel, D. Quéré, J. Rabe, H. Schäfer, J. Schreiber, M. Stamm, M. Steinhart, G. Subklew, J. Tomas, K. Vasilev, K. Wandelt, B. Wenclawiak, R. Wepf, R. Wiesendanger, D.Y. Yoon, M. Zharnikov, and U. Zimmermann.

*Hans-Jürgen Butt, Karlheinz Graf, and Michael Kappl*

Mainz, August 2003

# Contents

**Preface** **V**

**1 Introduction** **1**

**2 Liquid surfaces** **4**
2.1 Microscopic picture of the liquid surface . . . . . 4
2.2 Surface tension . . . . . 5
2.3 Equation of Young and Laplace . . . . . 8
2.3.1 Curved liquid surfaces . . . . . 8
2.3.2 Derivation of the Young–Laplace equation . . . . . 10
2.3.3 Applying the Young–Laplace equation . . . . . 11
2.4 Techniques to measure the surface tension . . . . . 12
2.5 The Kelvin equation . . . . . 15
2.6 Capillary condensation . . . . . 17
2.7 Nucleation theory . . . . . 20
2.8 Summary . . . . . 23
2.9 Exercises . . . . . 24

**3 Thermodynamics of interfaces** **26**
3.1 The surface excess . . . . . 26
3.2 Fundamental thermodynamic relations . . . . . 29
3.2.1 Internal energy and Helmholtz energy . . . . . 29
3.2.2 Equilibrium conditions . . . . . 30
3.2.3 Location of the interface . . . . . 31
3.2.4 Gibbs energy and definition of the surface tension . . . . . 32
3.2.5 Free surface energy, interfacial enthalpy and Gibbs surface energy . . 32
3.3 The surface tension of pure liquids . . . . . 34
3.4 Gibbs adsorption isotherm . . . . . 35
3.4.1 Derivation . . . . . 36
3.4.2 System of two components . . . . . 37
3.4.3 Experimental aspects . . . . . 38
3.4.4 The Marangoni effect . . . . . 39
3.5 Summary . . . . . 40
3.6 Exercises . . . . . 41

**4 The electric double layer** **42**
4.1 Introduction . . . 42
4.2 Poisson–Boltzmann theory of the diffuse double layer . . . 43
4.2.1 The Poisson–Boltzmann equation . . . 43
4.2.2 Planar surfaces . . . 44
4.2.3 The full one-dimensional case . . . 46
4.2.4 The Grahame equation . . . 49
4.2.5 Capacity of the diffuse electric double layer . . . 50
4.3 Beyond Poisson–Boltzmann theory . . . 50
4.3.1 Limitations of the Poisson–Boltzmann theory . . . 50
4.3.2 The Stern layer . . . 52
4.4 The Gibbs free energy of the electric double layer . . . 54
4.5 Summary . . . 55
4.6 Exercises . . . 56

**5 Effects at charged interfaces** **57**
5.1 Electrocapillarity . . . 57
5.1.1 Theory . . . 58
5.1.2 Measurement of electrocapillarity . . . 60
5.2 Examples of charged surfaces . . . 61
5.2.1 Mercury . . . 62
5.2.2 Silver iodide . . . 63
5.2.3 Oxides . . . 65
5.2.4 Mica . . . 66
5.2.5 Semiconductors . . . 67
5.3 Measuring surface charge densities . . . 68
5.3.1 Potentiometric colloid titration . . . 68
5.3.2 Capacitances . . . 71
5.4 Electrokinetic phenomena: The zeta potential . . . 72
5.4.1 The Navier–Stokes equation . . . 72
5.4.2 Electro-osmosis and streaming potential . . . 73
5.4.3 Electrophoresis and sedimentation potential . . . 76
5.5 Types of potentials . . . 77
5.6 Summary . . . 79
5.7 Exercises . . . 79

**6 Surface forces** **80**
6.1 Van der Waals forces between molecules . . . 80
6.2 The van der Waals force between macroscopic solids . . . 84
6.2.1 Microscopic approach . . . 84
6.2.2 Macroscopic calculation — Lifshitz theory . . . 87
6.2.3 Surface energy and Hamaker constant . . . 92
6.3 Concepts for the description of surface forces . . . 93
6.3.1 The Derjaguin approximation . . . 93
6.3.2 The disjoining pressure . . . 95

6.4 Measurement of surface forces . . . 96
6.5 The electrostatic double-layer force . . . 98
6.5.1 General equations . . . 98
6.5.2 Electrostatic interaction between two identical surfaces . . . 101
6.5.3 The DLVO theory . . . 102
6.6 Beyond DLVO theory . . . 104
6.6.1 The solvation force and confined liquids . . . 104
6.6.2 Non DLVO forces in an aqueous medium . . . 106
6.7 Steric interaction . . . 107
6.7.1 Properties of polymers . . . 107
6.7.2 Force between polymer coated surfaces . . . 108
6.8 Spherical particles in contact . . . 111
6.9 Summary . . . 115
6.10 Exercises . . . 116

**7 Contact angle phenomena and wetting 118**
7.1 Young's equation . . . 118
7.1.1 The contact angle . . . 118
7.1.2 Derivation . . . 119
7.1.3 The line tension . . . 121
7.1.4 Complete wetting . . . 121
7.2 Important wetting geometries . . . 122
7.2.1 Capillary rise . . . 122
7.2.2 Particles in the liquid–gas interface . . . 123
7.2.3 Network of fibres . . . 125
7.3 Measurement of the contact angle . . . 126
7.3.1 Experimental methods . . . 126
7.3.2 Hysteresis in contact angle measurements . . . 128
7.3.3 Surface roughness and heterogeneity . . . 129
7.4 Theoretical aspects of contact angle phenomena . . . 131
7.5 Dynamics of wetting and dewetting . . . 133
7.5.1 Wetting . . . 133
7.5.2 Dewetting . . . 137
7.6 Applications . . . 138
7.6.1 Flotation . . . 138
7.6.2 Detergency . . . 140
7.6.3 Microfluidics . . . 141
7.6.4 Adjustable wetting . . . 142
7.7 Summary . . . 144
7.8 Exercises . . . 144

**8 Solid surfaces 145**
8.1 Introduction . . . 145
8.2 Description of crystalline surfaces . . . 146
8.2.1 The substrate structure . . . 146

8.2.2 Surface relaxation and reconstruction . . . 147
8.2.3 Description of adsorbate structures . . . 149
8.3 Preparation of clean surfaces . . . 150
8.4 Thermodynamics of solid surfaces . . . 153
8.4.1 Surface stress and surface tension . . . 153
8.4.2 Determination of the surface energy . . . 154
8.4.3 Surface steps and defects . . . 157
8.5 Solid–solid boundaries . . . 159
8.6 Microscopy of solid surfaces . . . 162
8.6.1 Optical microscopy . . . 162
8.6.2 Electron microscopy . . . 162
8.6.3 Scanning probe microscopy . . . 164
8.7 Diffraction methods . . . 167
8.7.1 Diffraction patterns of two-dimensional periodic structures . . . 168
8.7.2 Diffraction with electrons, X-rays, and atoms . . . 168
8.8 Spectroscopic methods . . . 171
8.8.1 Spectroscopy using mainly inner electrons . . . 171
8.8.2 Spectroscopy with outer electrons . . . 173
8.8.3 Secondary ion mass spectrometry . . . 174
8.9 Summary . . . 175
8.10 Exercises . . . 176

**9 Adsorption 177**
9.1 Introduction . . . 177
9.1.1 Definitions . . . 177
9.1.2 The adsorption time . . . 178
9.1.3 Classification of adsorption isotherms . . . 179
9.1.4 Presentation of adsorption isotherms . . . 181
9.2 Thermodynamics of adsorption . . . 182
9.2.1 Heats of adsorption . . . 182
9.2.2 Differential quantities of adsorption and experimental results . . . 184
9.3 Adsorption models . . . 185
9.3.1 The Langmuir adsorption isotherm . . . 185
9.3.2 The Langmuir constant and the Gibbs energy of adsorption . . . 188
9.3.3 Langmuir adsorption with lateral interactions . . . 189
9.3.4 The BET adsorption isotherm . . . 189
9.3.5 Adsorption on heterogeneous surfaces . . . 192
9.3.6 The potential theory of Polanyi . . . 193
9.4 Experimental aspects of adsorption from the gas phase . . . 195
9.4.1 Measurement of adsorption isotherms . . . 195
9.4.2 Procedures to measure the specific surface area . . . 198
9.4.3 Adsorption on porous solids — hysteresis . . . 199
9.4.4 Special aspects of chemisorption . . . 201
9.5 Adsorption from solution . . . 202
9.6 Summary . . . 203
9.7 Exercises . . . 205

**10 Surface modification 206**
10.1 Introduction . . . 206
10.2 Chemical vapor deposition . . . 207
10.3 Soft matter deposition . . . 209
10.3.1 Self-assembled monolayers . . . 209
10.3.2 Physisorption of Polymers . . . 212
10.3.3 Polymerization on surfaces . . . 215
10.4 Etching techniques . . . 217
10.5 Summary . . . 221
10.6 Exercises . . . 221

**11 Friction, lubrication, and wear 223**
11.1 Friction . . . 223
11.1.1 Introduction . . . 223
11.1.2 Amontons' and Coulomb's Law . . . 224
11.1.3 Static, kinetic, and stick-slip friction . . . 226
11.1.4 Rolling friction . . . 228
11.1.5 Friction and adhesion . . . 229
11.1.6 Experimental Aspects . . . 230
11.1.7 Techniques to measure friction . . . 230
11.1.8 Macroscopic friction . . . 232
11.1.9 Microscopic friction . . . 232
11.2 Lubrication . . . 236
11.2.1 Hydrodynamic lubrication . . . 236
11.2.2 Boundary lubrication . . . 238
11.2.3 Thin film lubrication . . . 239
11.2.4 Lubricants . . . 240
11.3 Wear . . . 241
11.4 Summary . . . 244
11.5 Exercises . . . 245

**12 Surfactants, micelles, emulsions, and foams 246**
12.1 Surfactants . . . 246
12.2 Spherical micelles, cylinders, and bilayers . . . 250
12.2.1 The critical micelle concentration . . . 250
12.2.2 Influence of temperature . . . 252
12.2.3 Thermodynamics of micellization . . . 253
12.2.4 Structure of surfactant aggregates . . . 255
12.2.5 Biological membranes . . . 258
12.3 Macroemulsions . . . 259
12.3.1 General properties . . . 259
12.3.2 Formation . . . 261
12.3.3 Stabilization . . . 262
12.3.4 Evolution and aging . . . 265
12.3.5 Coalescence and demulsification . . . 267

12.4 Microemulsions . . . 268
12.4.1 Size of droplets . . . 268
12.4.2 Elastic properties of surfactant films . . . 269
12.4.3 Factors influencing the structure of microemulsions . . . 270
12.5 Foams . . . 272
12.5.1 Classification, application and formation . . . 272
12.5.2 Structure of foams . . . 274
12.5.3 Soap films . . . 274
12.5.4 Evolution of foams . . . 277
12.6 Summary . . . 278
12.7 Exercises . . . 279

**13 Thin films on surfaces of liquids 280**
13.1 Introduction . . . 280
13.2 Phases of monomolecular films . . . 283
13.3 Experimental techniques to study monolayers . . . 286
13.3.1 Optical methods . . . 286
13.3.2 X-ray reflection and diffraction . . . 287
13.3.3 The surface potential . . . 290
13.3.4 Surface elasticity and viscosity . . . 292
13.4 Langmuir–Blodgett transfer . . . 293
13.5 Thick films – spreading of one liquid on another . . . 295
13.6 Summary . . . 297
13.7 Exercises . . . 297

**14 Solutions to exercises 299**

**Appendix**

**A Analysis of diffraction patterns 321**
A.1 Diffraction at three dimensional crystals . . . 321
A.1.1 Bragg condition . . . 321
A.1.2 Laue condition . . . 321
A.1.3 The reciprocal lattice . . . 323
A.1.4 Ewald construction . . . 325
A.2 Diffraction at Surfaces . . . 325
A.3 Intensity of diffraction peaks . . . 327

**B Symbols and abbreviations 331**

**Bibliography 335**

**Index 355**

# 1 Introduction

An interface is the area which separates two phases from each other. If we consider the solid, liquid, and gas phase we immediately get three combinations of interfaces: the solid–liquid, the solid–gas, and the liquid–gas interface. These interfaces are also called surfaces. Interface is, however, a more general term than surface. Interfaces can also separate two immiscible liquids such as water and oil. These are called liquid–liquid interfaces. Solid–solid interfaces separate two solid phases. They are important for the mechanical behavior of solid materials. Gas–gas interfaces do not exist because gases mix.

Often interfaces and colloids are discussed together. Colloids are disperse systems, in which one phase has dimensions in the order of 1 nm to 1 $\mu$m (see Fig. 1.1). The word "colloid" comes from the Greek word for glue and has been used the first time in 1861 by Graham[1]. He applied it to materials which seemed to dissolve but were not able to penetrate a membrane, such as albumin, starch, and dextrin. A dispersion is a two-phase system which is uniform on the macroscopic but not on the microscopic scale. It consists of grains or droplets of one phase in a matrix of the other phase.

Different kinds of dispersions can be formed. Most of them have important applications and have special names (Table 1.1). While there are only five types of interface, we can distinguish ten types of disperse system because we have to discriminate between the continuous, dispersing (external) phase and the dispersed (inner) phase. In some cases this distinction is obvious. Nobody will, for instance, mix up fog with a foam although in both cases a liquid and a gas are involved. In other cases the distinction between continuous and inner phase cannot be made because both phases might form connected networks. Some emulsions for instance tend to form a bicontinuous phase, in which both phases form an interwoven network.

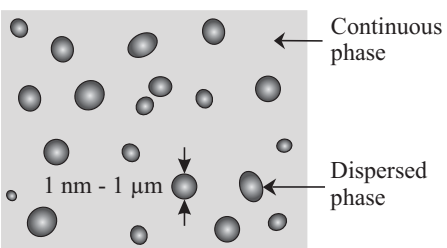



**Figure 1.1:** Schematic of a dispersion.

Colloids and interfaces are intimately related. This is a direct consequence of their enormous specific surface area. More precisely: their interface-to-volume relation is so large, that their behavior is completely determined by surface properties. Gravity is negligible in the

[1] Thomas Graham, 1805–1869. British chemist, professor in Glasgow and London.

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

**Table 1.1:** Types of dispersions. *Porous solids have a bicontinuous structure while in a solid foam the gas phase is clearly dispersed.

| Continuous phase | Dispersed phase | Term | Example |
|---|---|---|---|
| Gas | liquid | aerosol | clouds, fog, smog, hairspray |
| | solid | aerosol | smoke, dust, pollen |
| Liquid | gas | foam | lather, whipped cream, foam on beer |
| | liquid | emulsion | milk |
| | solid | sol | ink, muddy water, dispersion paint |
| Solid | gas | porous solids* | |
| | | foam | styrofoam, soufflés |
| | liquid | solid emulsion | butter |
| | solid | solid suspension | concrete |

majority of cases. For this reason we could define colloidal systems as systems which are dominated by interfacial effects rather than bulk properties. This is also the reason why interfacial science is the basis for nanoscience and technology and many inventions in this new field originate from surface science.

■ **Example 1.1.** A system which is dominated by surface effects is shown on the left side of the cover. The scanning electronic microscope (SEM) image shows aggregates of quartz ($SiO_2$) particles (diameter 0.9 $\mu$m). These particles were blown as dust into a chamber filled with gas. While sedimenting they formed fractal aggregates due to attractive van der Waals forces. On the bottom they were collected. These aggregates are stable for weeks and months and even shaking does not change their structure. Gravity and inertia, which rule the macroscopic world, are not able to bend down the particle chains. Surface forces are much stronger.

In the recent literature the terms nanoparticles and nanosystems are used, in analogy to colloid and colloidal systems. The prefix "nano" indicates dimensions in the 1 to 100 nm range. This is above the atomic scale and, unless highly refined methods are used, below the resolution of a light microscope and thus also below the accuracy of optical microstructuring techniques.

Why is there an interest in interfaces and colloids? First, for a better understanding of natural processes. For example, in biology the surface tension of water allows to form lipid membranes. This is a prerequisite for the formation of compartments and thus any form of life. In geology the swelling of clay or soil in the presence of water is an important process. The formation of clouds and rain due to nucleation of water around small dust particles is dominated by surface effects. Many foods, like butter, milk, or mayonnaise are emulsions. Their properties are determined by the liquid–liquid interface. Second, there are many technological applications. One such example is flotation in mineral processing or the bleaching of scrap paper. Washing and detergency are examples which any person encounters every day.

Often the production of new materials such as composite materials heavily involves processes at interfaces. Thin films on surfaces are often dominated by surface effects. Examples are latex-films, coatings, and paints. The flow behavior of powders and granular media is determined by surface forces. In tribology, wear is reduced by lubrication which again is a surface phenomenon.

Typical for many of the industrial applications is a very refined and highly developed technology, but only a limited understanding of the underlying fundamental processes. A better understanding is, however, required to further improve the efficiency or reduce dangers to the environment.

Introductory books on interface science are Refs. [1–6]. For a deeper understanding we recommend the series of books of Lyklema [7–9].

# 2 Liquid surfaces

## 2.1 Microscopic picture of the liquid surface

A surface is not an infinitesimal sharp boundary in the direction of its normal, but it has a certain thickness. For example, if we consider the density $\rho$ normal to the surface (Fig. 2.1), we can observe that, within a few molecules, the density decreases from that of the bulk liquid to that of its vapor [10].

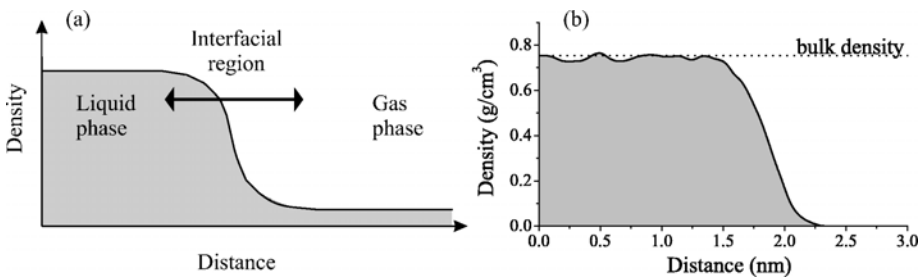



**Figure 2.1:** Density of a liquid versus the coordinate normal to its surface: (a) is a schematic plot; (b) results from molecular dynamics simulations of a $n$-tridecane ($C_{13}H_{28}$) at 27°C adapted from Ref. [11]. Tridecane is practically not volatile. For this reason the density in the vapor phase is negligible.

The density is only one criterion to define the thickness of an interface. Another possible parameter is the orientation of the molecules. For example, water molecules at the surface prefer to be oriented with their negative sides "out" towards the vapor phase. This orientation fades with increasing distance from the surface. At a distance of 1–2 nm the molecules are again randomly oriented.

Which thickness do we have to use? This depends on the relevant parameter. If we are for instance, interested in the density of a water surface, a realistic thickness is in the order of 1 nm. Let us assume that a salt is dissolved in the water. Then the concentration of ions might vary over a much larger distance (characterized by the Debye length, see Section 4.2.2). With respect to the ion concentration, the thickness is thus much larger. In case of doubt, it is safer to choose a large value for the thickness.

The surface of a liquid is a very turbulent place. Molecules evaporate from the liquid into the vapor phase and vice versa. In addition, they diffuse into the bulk phase and molecules from the bulk diffuse to the surface.

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

■ **Example 2.1.** To estimate the number of gas molecules hitting the liquid surface per second, we recall the kinetic theory of ideal gases. In textbooks of physical chemistry the rate of effusion of an ideal gas through a small hole is given by [12]

$$\frac{PA}{\sqrt{2\pi m k_B T}} \tag{2.1}$$

Here, $A$ is the cross-sectional area of the hole and $m$ is the molecular mass. This is equal to the number of water molecules hitting a surface area $A$ per second. Water at 25°C has a vapor pressure $P$ of 3168 Pa. With a molecular mass $m$ of $0.018\ \mathrm{kgmol}^{-1}/6.02 \times 10^{23}\,\mathrm{mol}^{-1} \approx 3 \times 10^{-26}\,\mathrm{kg}$, $10^7$ water molecules per second hit a surface area of 10 Å$^2$. In equilibrium the same number of molecules escape from the liquid phase. 10 Å$^2$ is approximately the area covered by one water molecule. Thus, the average time a water molecule remains on the surface is of the order of 0.1 $\mu$s.

## 2.2 Surface tension

The following experiment helps us to define the most fundamental quantity in surface science: the surface tension. A liquid film is spanned over a frame, which has a mobile slider (Fig. 2.2). The film is relatively thick, say $1\mu$m, so that the distance between the back and front surfaces is large enough to avoid overlapping of the two interfacial regions. Practically, this experiment might be tricky even in the absence of gravity but it does not violate a physical law so that it is in principle feasible. If we increase the surface area by moving the slider a distance $dx$ to the right, work has to be done. This work $dW$ is proportional to the increase in surface area $dA$. The surface area increases by twice $b \cdot dx$ because the film has a front and back side. Introducing the proportionality constant $\gamma$ we get

$$dW = \gamma \cdot dA \tag{2.2}$$

The constant $\gamma$ is called surface tension.

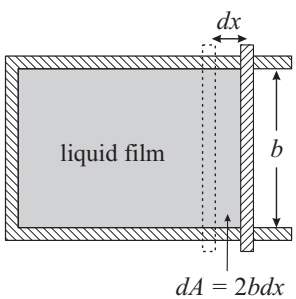



**Figure 2.2:** Schematic set-up to verify Eq. (2.2) and define the surface tension.

Equation (2.2) is an empirical law and a definition at the same time. The empirical law is that the work is proportional to the change in surface area. This is not only true for infinitesimal small changes of $A$ (which is trivial) but also for significant increases of the surface area: $\Delta W = \gamma \cdot \Delta A$. In general, the proportionality constant depends on the composition of the liquid and the vapor, temperature, and pressure, but it is independent of the area. The definition is that we call the proportionality constant "surface tension".

The surface tension can also be defined by the force $F$ that is required to hold the slider in place and to balance the surface tensional force:

$$|F| = 2\gamma b \tag{2.3}$$

Both forms of the law are equivalent, provided that the process is reversible. Then we can write

$$F = -\frac{dW}{dx} = -2\gamma b \tag{2.4}$$

The force is directed to the left while $x$ increases to the right. Therefore we have a negative sign.

The unit of surface tension is either J/m$^2$ or N/m. Surface tensions of liquids are of the order of 0.02–0.08 N/m (Table 2.1). For convenience they are usually given in mN/m (or $10^{-3}$ N/m), where the first "m" stands for "milli".

The term "surface tension" is tied to the concept that the surface stays under a tension. In a way, this is similar to a rubber balloon, where also a force is required to increase the surface area of its rubber membrane against a tension. There is, however, a difference: while the expansion of a liquid surface is a plastic process the stretching of a rubber membrane is usually elastic.

**Table 2.1:** Surface tensions $\gamma$ of some liquids at different temperatures $T$.

| Substance | $T$ | $\gamma$ ( mNm$^{-1}$) | Substance | $T$ | $\gamma$ (mNm$^{-1}$) |
|---|---|---|---|---|---|
| Water | 10°C | 74.23 | Mercury | 25°C | 485.48 |
| | 25°C | 71.99 | Phenol | 50°C | 38.20 |
| | 50°C | 67.94 | Benzene | 25°C | 28.22 |
| | 75°C | 63.57 | Toluene | 25°C | 27.93 |
| | 100°C | 58.91 | Dichloromethane | 25°C | 27.20 |
| Argon | 90 K | 11.90 | $n$-pentane | 25°C | 15.49 |
| Methanol | 25°C | 22.07 | $n$-hexane | 25°C | 17.89 |
| Ethanol | 10°C | 23.22 | $n$-heptane | 25°C | 19.65 |
| | 25°C | 21.97 | $n$-octane | 10°C | 22.57 |
| | 50°C | 19.89 | | 25°C | 21.14 |
| 1-propanol | 25°C | 23.32 | | 50°C | 18.77 |
| 1-butanol | 25°C | 24.93 | | 75°C | 16.39 |
| 2-butanol | 25°C | 22.54 | | 100°C | 14.01 |
| Acetone | 25°C | 23.46 | Formamide | 25°C | 57.03 |

■ **Example 2.2.** If a water film is formed on a frame with a slider length of 1 cm, then the film pulls on the slider with a force of

$$2 \times 0.01 \text{ m} \times 0.072 \text{ Jm}^{-2} = 1.44 \times 10^{-3}\text{N}$$

That corresponds to a weight of 0.15 g.

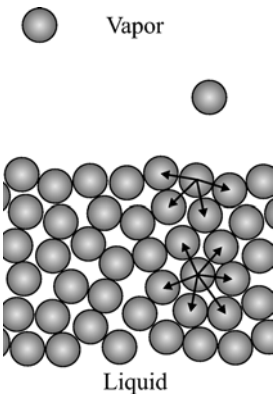



**Figure 2.3:** Schematic molecular structure of a liquid–vapor interface.

How can we interpret the concept of surface tension on the molecular level? For molecules it is energetically favorable to be surrounded by other molecules. Molecules attract each other by different interactions such as van der Waals forces or hydrogen bonds (for details see Chapter 6). Without this attraction there would not be a condensed phase at all, there would only be a vapor phase. The sheer existence of a condensed phase is evidence for an attractive interaction between the molecules. At the surface, molecules are only partially surrounded by other molecules and the number of adjacent molecules is smaller than in the bulk (Fig. 2.3). This is energetically unfavorable. In order to bring a molecule from the bulk to the surface, work has to be done. With this view $\gamma$ can be interpreted as the energy required to bring molecules from inside the liquid to the surface and to create new surface area. Therefore often the term "surface energy" is used for $\gamma$. As we shall see in the next chapter this might lead to some confusion. To avoid this we use the term surface tension.

With this interpretation of the surface tension in mind we immediately realize that $\gamma$ has to be positive. Otherwise the Gibbs free energy of interaction would be repulsive and all molecules would immediately evaporate into the gas phase.

■ **Example 2.3.** Estimate the surface tension of cyclohexane from the energy of vaporization $\Delta_{vap}U = 30.5$ kJ/mol at 25°C. The density of cyclohexane is $\rho = 773$ kg/m$^3$, its molecular weight is $M = 84.16$ g/mol.

For a rough estimate we picture the liquid as being arranged in a cubic structure. Each molecule is surrounded by 6 nearest neighbors. Thus each bond contributes roughly $\Delta_{vap}U/6 = 5.08$ kJ/mol. At the surface one neighbor and hence one bond is missing. Per mole we therefore estimate a surface tension of 5.08 kJ/mol.

To estimate the surface tension we need to know the surface area occupied by one molecule. If the molecules form a cubic structure, the volume of one unit cell is $a^3$, where $a$ is the distance between nearest neighbors. This distance can be calculated from the density:

$$\begin{aligned} a^3 &= \frac{M}{\rho N_A} = \frac{0.08416\ \mathrm{kg/mol}}{773\ \mathrm{kg/m^3} \cdot 6.02 \times 10^{23}\ \mathrm{mol^{-1}}} = 1.81 \times 10^{-28}\ \mathrm{m^3} \Rightarrow \\ a &= 0.565\ \mathrm{nm} \end{aligned}$$

The surface area per molecule is $a^2$. For the surface energy we estimate

$$\gamma = \frac{\Delta_{vap}U}{6N_A a^2} = \frac{5080\ \mathrm{Jmol}^{-1}}{6.02 \times 10^{23}\,\mathrm{mol}^{-1} \cdot (0.565 \times 10^{-9}\,\mathrm{m})^2} = 0.0264\,\frac{\mathrm{J}}{\mathrm{m}^2}$$

For such a rough estimate the result is surprisingly close to the experimental value of 0.0247 J/m$^2$.

## 2.3 Equation of Young and Laplace

### 2.3.1 Curved liquid surfaces

We start by describing an important phenomenon: If in equilibrium a liquid surface is curved, there is a pressure difference across it. To illustrate this let us consider a circular part of the surface. The surface tension tends to minimize the area. This results in a planar geometry of the surface. In order to curve the surface, the pressure on one side must be larger than on the other side. The situation is much like that of a rubber membrane. If we, for instance, take a tube and close one end with a rubber membrane, the membrane will be planar (provided the membrane is under some tension) (Fig. 2.4). It will remain planar as long as the tube is open at the other end and the pressure inside the tube is equal to the outside pressure. If we now blow carefully into the tube, the membrane bulges out and becomes curved due to the increased pressure inside the tube. If we suck on the tube, the membrane bulges inside the tube because now the outside pressure is higher than the pressure inside the tube.

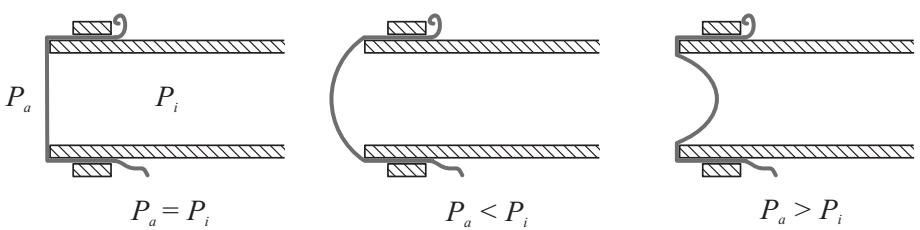



**Figure 2.4:** Rubber membrane at the end of a cylindrical tube. An inner pressure $P_i$ can be applied, which is different than the outside pressure $P_a$.

The Young[1]–Laplace[2] equation relates the pressure difference between the two phases $\Delta P$ and the curvature of the surface:

$$\Delta P = \gamma \cdot \left( \frac{1}{R_1} + \frac{1}{R_2} \right) \tag{2.5}$$

$R_1$ and $R_2$ are the two principal radii of curvature. $\Delta P$ is also called Laplace pressure. Equation (2.5) is also referred to as the Laplace equation.

[1] Thomas Young, 1773–1829. English physician and physicist, professor in Cambridge.

[2] Pierre-Simon Laplace, Marquis de Laplace, 1749–1827. French natural scientist.

It is perhaps worthwhile to describe the principal radii of curvature in a little bit more detail. The curvature $1/R_1 + 1/R_2$ at a point on an arbitrarily curved surface is obtained as follows. At the point of interest we draw a normal through the surface and then pass a plane through this line and the intersection of this line with the surface. The line of intersection will, in general, be curved at the point of interest. The radius of curvature $R_1$ is the radius of a circle inscribed to the intersection at the point of interest. The second radius of curvature is obtained by passing a second plane through the surface also containing the normal, but perpendicular to the first plane. This gives the second intersection and leads to the second radius of curvature $R_2$. So the planes defining the radii of curvature must be perpendicular to each other and contain the surface normal. Otherwise their orientation is arbitrary. A law of differential geometry says that the value $1/R_1 + 1/R_2$ for an arbitrary surface does not depend on the orientation, as long as the radii are determined in perpendicular directions.

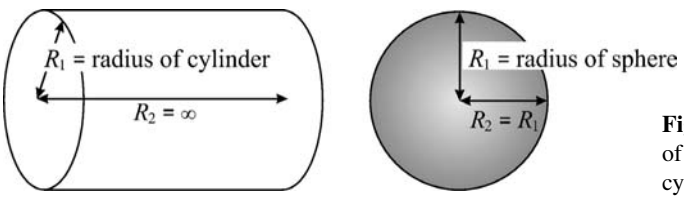



**Figure 2.5:** Illustration of the curvature of a cylinder and a sphere.

Let us illustrate the curvature for two examples. For a cylinder of radius $r$ a convenient choice is $R_1 = r$ and $R_2 = \infty$ so that the curvature is $1/r + 1/\infty = 1/r$. For a sphere with radius $R$ we have $R_1 = R_2$ and the curvature is $1/R + 1/R = 2/R$ (Fig. 2.5).

■ **Example 2.4.** How large is the pressure in a spherical bubble with a diameter of 2 mm and a bubble of 20 nm diameter in pure water, compared with the pressure outside? For a bubble the curvature is identical to that of a sphere: $R_1 = R_2 = R$. Therefore

$$\Delta P = \frac{2\gamma}{R} \tag{2.6}$$

With $R = 1$ mm we get

$$\Delta P = 0.072\ \frac{\mathrm{J}}{\mathrm{m}^2} \times \frac{2}{10^{-3}\mathrm{m}} = 144\ \mathrm{Pa}$$

With $R = 10$ nm the pressure is $\Delta P = 0.072\ \mathrm{J/m}^2 \times 2/10^{-8}\,\mathrm{m} = 1.44 \times 10^7$ Pa $=$ 144 bar. The pressure inside the bubbles is therefore 144 Pa and $1.44 \times 10^7$ Pa, respectively, higher than the outside pressure.

The Young–Laplace equation has several fundamental implications:

- If we know the shape of a liquid surface we know its curvature and we can calculate the pressure difference.
- In the absence of external fields (e.g. gravity), the pressure is the same everywhere in the liquid; otherwise there would be a flow of liquid to regions of low pressure. Thus, $\Delta P$ is constant and Young–Laplace equation tells us that in this case the surface of the liquid has the same curvature everywhere.

- With the help of the Young–Laplace Eq. (2.5) it is possible to calculate the equilibrium shape of a liquid surface. If we know the pressure difference and some boundary conditions (such as the volume of the liquid and its contact line) we can calculate the geometry of the liquid surface.

In practice, it is usually not trivial to calculate the geometry of a liquid surface with Eq. (2.5). The shape of the liquid surface can mathematically be described by a function $z = z(x, y)$. The $z$ coordinate of the surface is given as a function of its $x$ and $y$ coordinate. The curvature involves the second derivative. As a result, calculating the shape of a liquid surface involves solving a partial differential equation of second order, which is certainly not a simple task.

In many cases we deal with rotational symmetric structures. Assuming that the axis of symmetry is identical to the $y$ axis of an orthogonal cartesian coordinate system, then it is convenient to put one radius of curvature in the plane of the *xy* coordinate. This radius is given by

$$\frac{1}{R_1} = \frac{y''}{\sqrt{(1 + y'^2)^3}}, \tag{2.7}$$

where $y'$ and $y''$ are the first and second derivatives with respect to $x$. The plane for the second bending radius is perpendicular to the *xy* plane. It is

$$\frac{1}{R_2} = \frac{y'}{x\sqrt{1 + y'^2}} \tag{2.8}$$

### 2.3.2 Derivation of the Young–Laplace equation

To derive the equation of Young and Laplace we consider a small part of a liquid surface. First, we pick a point X and draw a line around it which is characterized by the fact that all points on that line are the same distance $d$ away from X (Fig. 2.6). If the liquid surface is planar, this would be a flat circle. On this line we take two cuts that are perpendicular to each other (AXB and CXD). Consider in B a small segment on the line of length $dl$. The surface tension pulls with a force $\gamma\, dl$. The vertical force on that segment is $\gamma\, dl\, \sin\alpha$. For small surface areas (and small $\alpha$) we have $\sin\alpha \approx d/R_1$ where $R_1$ is the radius of curvature along AXB. The vertical force component is

$$\gamma \cdot dl \cdot \frac{d}{R_1} \tag{2.9}$$

The sum of the four vertical components at points A, B, C, and D is

$$\gamma \cdot dl \cdot \left(\frac{2d}{R_1} + \frac{2d}{R_2}\right) = \gamma \cdot dl \cdot 2d \cdot \left(\frac{1}{R_1} + \frac{1}{R_2}\right) \tag{2.10}$$

This expression is independent of the absolute orientation of AB and CD. Integration over the borderline (only 90° rotation of the four segments) gives the total vertical force, caused by the surface tension:

$$\pi d^2 \cdot \gamma \cdot \left(\frac{1}{R_1} + \frac{1}{R_2}\right) \tag{2.11}$$

In equilibrium, this downward force must be compensated by an equal force in the opposite direction. This upward force is caused by an increased pressure $\Delta P$ on the concave side of $\pi d^2 \Delta P$. Equating both forces leads to

$$\Delta P \cdot \pi d^2 = \pi d^2 \cdot \gamma \cdot \left( \frac{1}{R_1} + \frac{1}{R_2} \right) \Rightarrow \Delta P = \gamma \cdot \left( \frac{1}{R_1} + \frac{1}{R_2} \right) \tag{2.12}$$

These considerations are valid for any small part of the liquid surface. Since the part is arbitrary the Young–Laplace equation must be valid everywhere.

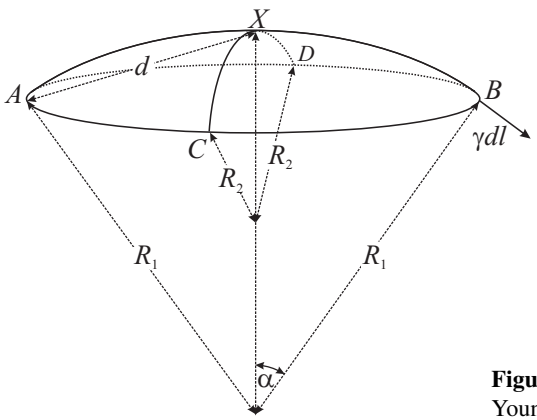



**Figure 2.6:** Diagram used for deriving the Young–Laplace equation.

### 2.3.3 Applying the Young–Laplace equation

When applying the equation of Young and Laplace to simple geometries it is usually obvious at which side the pressure is higher. For example, both inside a bubble and inside a drop, the pressure is higher than outside (Fig. 2.7). In other cases this is not so obvious because the curvature can have an opposite sign. One example is a drop hanging between the planar ends of two cylinders (Fig. 2.7). Then the two principal curvatures, defined by

$$C_1 = \frac{1}{R_1} \qquad \text{and} \qquad C_2 = \frac{1}{R_2}, \tag{2.13}$$

can have a different sign. We count it positive if the interface is curved towards the liquid. The pressure difference is defined as $\Delta P = P_{liquid} - P_{gas}$.

■ **Example 2.5.** For a drop in a gaseous environment, the two principal curvatures are positive and given by $C_1 = C_2 = 1/R$. The pressure difference is positive, which implies that the pressure inside the liquid is higher than outside.

For a bubble in a liquid environment the two principal curvatures are negative: $C_1 = C_2 = -1/R$. The pressure difference is negative and the pressure inside the liquid is lower than inside the bubble.

For a drop hanging between the ends of two cylinders (Fig. 2.7B) in a gaseous environment, one curvature is conveniently chosen to be $C_1 = 1/R_1$. The other curvature is negative, $C_2 = -1/R_2$. The pressure difference depends on the specific values of $R_1$ and $R_2$.

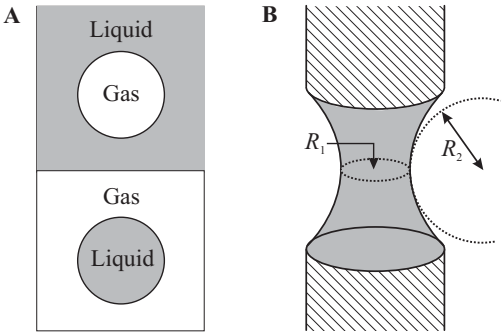



**Figure 2.7:** (A) A gas bubble in liquid and a drop in a gaseous environment. (B) A liquid meniscus with radii of curvature of opposite sign between two solid cylinders.

The shape of a liquid surface is determined by the Young–Laplace equation. In large structures we have to consider also the hydrostatic pressure. Then the equation of Young and Laplace becomes

$$\Delta P = \gamma \cdot \left( \frac{1}{R_1} + \frac{1}{R_2} \right) + \rho g h \tag{2.14}$$

Here, $g$ is the acceleration of free fall and $h$ is the height coordinate.

What is a large and what is a small structure? In practice this is a relevant question because for small structures we can neglect $\rho g h$ and use the simpler equation. Several authors define the capillary constant $\sqrt{2\gamma/\rho g}$ (as a source of confusion other authors have defined $\sqrt{\gamma/\rho g}$ as the capillary constant). For liquid structures whose curvature is much smaller than the capillary constant the influence of gravitation can be neglected. At 25°C the capillary constant is 3.8 mm for water and 2.4 mm for hexane.

## 2.4 Techniques to measure the surface tension

Before we can discuss the experimental techniques used to measure the surface tension, we need to introduce the so called contact angle $\Theta$. When we put a drop of liquid on a solid surface the edge usually forms a defined angle which depends only on the material properties of the liquid and the solid (Fig. 2.8). This is the contact angle. Here we only need to know what it is. In Chapter 8, contact angle phenomena are discussed in more detail. For a wetting surface we have $\Theta = 0$.

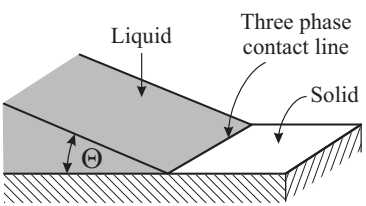



**Figure 2.8:** Rim of a liquid drop on a planar solid surface with its contact angle $\Theta$.

There are several techniques used to measure the surface tension of liquids. The most common technique is to measure optically the contour of a **sessile** or **pendant drop**. The measured contour is then fitted with a contour calculated using the Young–Laplace Eq. (2.5).

From this fit the surface tension is obtained. The same method is applied with a **pendant** or **sessile bubble**. Using a bubble ensures that the vapor pressure is 100%, a requirement for doing experiments in thermodynamic equilibrium. Often problems caused by contamination are reduced.

In the **Maximum-bubble-pressure method** the surface tension is determined from the value of the pressure which is necessary to push a bubble out of a capillary against the Laplace pressure. Therefore a capillary tube, with inner radius $r_C$, is immersed into the liquid (Fig. 2.9). A gas is pressed through the tube, so that a bubble is formed at its end. If the pressure in the bubble increases, the bubble is pushed out of the capillary more and more. In that way, the curvature of the gas–liquid interface increases according to the Young–Laplace equation. The maximum pressure is reached when the bubble forms a half-sphere with a radius $r_B = r_C$. This maximum pressure is related to the surface tension by $\gamma = r_C \Delta P/2$. If the volume of the bubble is further increased, the radius of the bubble would also have to become larger. A larger radius corresponds to a smaller pressure. The bubble would thus become unstable and detach from the capillary tube.

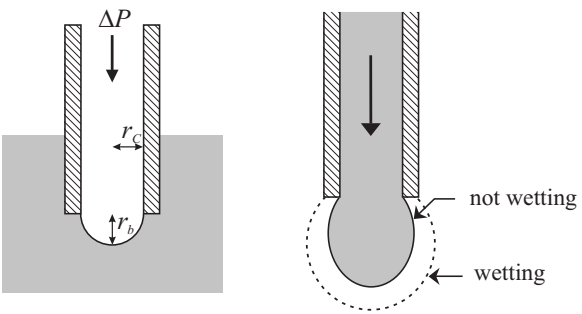



**Figure 2.9:** Maximal bubble pressure and drop-weight method to measure the surface tension of liquids.

**Drop-weight method**. Here, the liquid is allowed to flow out from the bottom of a capillary tube. Drops are formed which detach when they reach a critical dimension. The weight of a drop falling out of a capillary is measured. To get a precise measure, this is done for a number of drops and the total weight is divided by this number.

As long as the drop is still hanging at the end of the capillary, its weight is more than balanced by the surface tension. A drop falls off when the gravitational force $mg$, determined by the mass $m$ of the drop, is no longer balanced by the surface tension. The surface tensional force is equal to the surface tension multiplied by the circumference. This leads to

$$mg = 2\pi r_C \gamma \tag{2.15}$$

Thus, the mass is determined by the radius of the capillary. Here, we have to distinguish between the inner and outer diameter of the capillary. If the material of which the capillary is formed is not wetted by the liquid, the inner diameter enters into Eq. (2.15). If the surface of the capillary tube is wetted by the liquid, the external radius of the capillary has to be taken. For completely nonwetting surfaces (contact angle higher than 90°) the internal radius determines the drop weight. Experimentally, Tate already observed in 1864 that "other things being the same, the weight of a drop of a liquid is proportional to the diameter of the tube in which it is formed" [13].

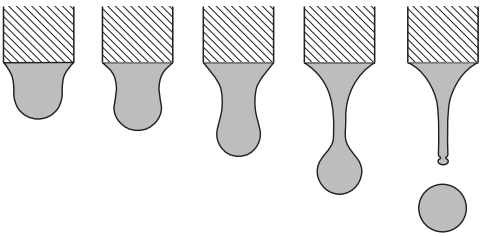


**Figure 2.10:** Release of a liquid drop from a capillary.

In practice, the equation is only approximately valid, and a weight less than the ideal value is measured. The reason becomes evident when the process of drop formation is observed closely (Fig. 2.10): A thin neck is formed before the drop is released. Correction factors $f$ are therefore used and Eq. (2.15) becomes: $mg = 2\pi f r_C \gamma$.

We have to admit that the maximum bubble pressure and the drop-weight methods are not very common for measuring the surface tension. Nevertheless we described them because the underlying phenomena, that is bubbling a gas into liquid and pressing a liquid out of a capillary, are very important. A common device used to measure $\gamma$ is the **ring tensiometer**, called also the **Du-Noüy**[3] tensiometer [14]. In a ring tensiometer the force necessary to detach a ring from the surface of a liquid is measured (Fig. 2.11). The force required for the detachment is

$$2\pi \cdot (r_i + r_a) \cdot \gamma \tag{2.16}$$

A necessary condition is that the ring surface must be completely wetting. A platinum wire is often used which can be annealed for cleaning before the measurement. Even in the early measurements it turned out that Eq. (2.16) was generally in serious error and that an empirical correction function is required [15].

A widely used technique is the **Wilhelmy** [4]**-plate method**. A thin plate of glass, platinum, or filter paper is vertically placed halfway into the liquid. In fact, the specific material is not important, as long as it is wetted by the liquid. Close to the three-phase contact line the liquid surface is oriented almost vertically (provided the contact angle is 0°). Thus the surface tension can exert a downward force. One measures the force required to prevent the plate from being drawn into the liquid. After subtracting the gravitational force this force is $2l\gamma$, where $l$ is the length of the plate. In honor of Ludwig Wilhelmy, who studied the force on a plate in detail, the method was named after him [16]. The Wilhelmy-plate method is simple and no correction factors are required. Care has to be taken to keep the plates clean and prevent contamination in air.

Finally there are dynamic methods to measure the surface tension. For example, a liquid jet is pushed out from a nozzle, which has an elliptic cross-section. The relaxation to a circular cross-section is observed. An advantage of this method is that we can measure changes of the surface tension, which might be caused by diffusion of amphiphilic substances to the surface.

---

[3] Pierre Lecomte du Noüy, 1883–1974. French scientist who worked in New York and Paris.

[4] Ludwig Ferdinand Wilhelmy, 1812–1864. German physicochemist.

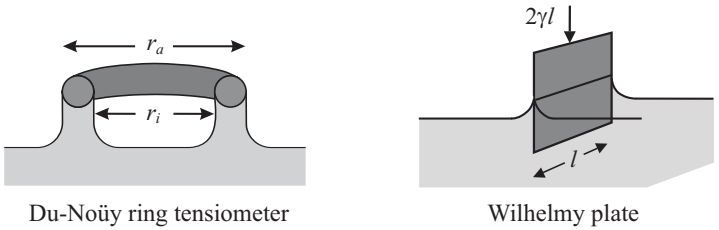



**Figure 2.11:** Du-Noüy ring tensiometer and Wilhelmy-plate method.

## 2.5 The Kelvin equation

In this chapter we get to know the second essential equation of surface science — the Kelvin[5] equation. Like the Young–Laplace equation it is based on thermodynamic principles and does not refer to a special material or special conditions. The subject of the Kelvin equation is the vapor pressure of a liquid. Tables of vapor pressures for various liquids and different temperatures can be found in common textbooks or handbooks of physical chemistry. These vapor pressures are reported for vapors which are in thermodynamic equilibrium with liquids having planar surfaces. When the liquid surface is curved, the vapor pressure changes. The vapor pressure of a drop is higher than that of a flat, planar surface. In a bubble the vapor pressure is reduced. The Kelvin equation tells us how the vapor pressure depends on the curvature of the liquid.

The cause for this change in vapor pressure is the Laplace pressure . The raised Laplace pressure in a drop causes the molecules to evaporate more easily. In the liquid, which surrounds a bubble, the pressure with respect to the inner part of the bubble is reduced. This makes it more difficult for molecules to evaporate. Quantitatively the change of vapor pressure for curved liquid surfaces is described by the Kelvin equation:

$$RT \cdot \ln \frac{P_0^{\mathrm{K}}}{P_0} = \gamma V_m \cdot \left( \frac{1}{R_1} + \frac{1}{R_2} \right) \tag{2.17}$$

$P_0^K$ is the vapor pressure of the curved, $P_0$ that of the flat surface. The index "0" indicates that everything is only valid in thermodynamic equilibrium. Please keep in mind: in equilibrium the curvature of a liquid surface is constant everywhere. $V_m$ is the molar volume of the liquid. For a sphere-like volume of radius $r$, the Kelvin equation can be simplified:

$$RT \cdot \ln \frac{P_0^K}{P_0} = \frac{2\gamma V_m}{r} \qquad \text{or} \qquad P_0^K = P_0 \cdot e^{\frac{2\gamma V_m}{RTr}} \tag{2.18}$$

The constant $2\gamma V_m/RT$ is 1.03 nm for Ethanol ($\gamma = 0.022$ N/m, $V_m = 58$ cm$^3$/mol) and 1.05 nm for Water ($\gamma = 0.072$ N/m, $V_m = 18$ cm$^3$/mol) at 25°C.

To derive the Kelvin equation we consider the Gibbs free energy of the liquid. The molar Gibbs free energy changes when the surface is being curved, because the pressure increases

[5] William Thomson, later Lord Kelvin, 1824–1907. Physics professor at the University of Glasgow.

due to the Laplace pressure. In general, any change in the Gibbs free energy is given by the fundamental equation $dG = VdP - SdT$. The increase of the Gibbs free energy per mole of liquid, upon curving, at constant temperature is

$$\Delta G_m = \int_0^{\Delta P} V_m dP = \gamma V_m \cdot \left( \frac{1}{R_1} + \frac{1}{R_2} \right), \tag{2.19}$$

We have assumed that the molar volume remains constant, which is certainly a reasonable assumption because most liquids are practically incompressible for the pressures considered. For a spherical drop in its vapor, we simply have $\Delta G_m = 2\gamma V_m/r$. The molar Gibbs free energy of the vapor depends on the vapor pressure $P_0$ according to

$$G_m = G_m^0 + RT \cdot \ln P_0 \tag{2.20}$$

For a liquid with a curved surface we have

$$G_m^K = G_m^0 + RT \cdot \ln P_0^K \tag{2.21}$$

The change of the molar Gibbs free energy inside the vapor due to curving the interface is therefore

$$\Delta G_m = G_m^K - G_m = RT \cdot \ln \frac{P_0^K}{P_0} \tag{2.22}$$

Since the liquid and vapor are supposed to be in equilibrium, the two expressions must be equal. This immediately leads to the Kelvin equation.

When applying the Kelvin equation, it is instructive to distinguish two cases: A drop in its vapor (or more generally: a positively curved liquid surface) and a bubble in liquid (a negatively curved liquid surface).

**Drop in its vapor**: The vapor pressure of a drop is higher than that of a liquid with a planar surface. One consequence is that an aerosol of drops (fog) should be unstable. To see this, let us assume that we have a box filled with many drops in a gaseous environment. Some drops are larger than others. The small drops have a higher vapor pressure than the large drops. Hence, more liquid evaporates from their surface. This tends to condense into large drops. Within a population of drops of different sizes, the bigger drops will grow at the expense of the smaller ones — a process called Ostwald ripening[6]. These drops will sink down and, at the end, bulk liquid fills the bottom of the box.

For a given vapor pressure, there is a critical drop size. Every drop bigger than this size will grow. Drops at a smaller size will evaporate. If a vapor is cooled to reach over-saturation, it cannot condense (because every drop would instantly evaporate again), unless nucleation sites are present. In that way it is possible to explain the existence of over-saturated vapors and also the undeniable existence of fog.

**Bubble in a liquid**: From Eq. (2.19) we see that a negative sign has to be used for a bubble because of the negative curvature of the liquid surface. As a result we get

$$RT \cdot \ln \frac{P_0^K}{P_0} = -\frac{2\gamma V_m}{r} \tag{2.23}$$

[6] In general, Ostwald ripening is the growth of long objects at the expense of smaller ones. Wilhelm Ostwald, 1853–1932. German physicochemist, professor in Leipzig, Nobel price for chemistry 1909.

Here, $r$ is the radius of the bubble. The vapor pressure inside a bubble is therefore reduced. This explains why it is possible to overheat liquids: When the temperature is increased above the boiling point (at a given external pressure) occasionally, tiny bubbles are formed. Inside the bubble the vapor pressure is reduced, the vapor condenses, and the bubble collapses. Only if a bubble larger than a certain critical size is formed, is it more likely to increases in size rather than to collapse. As an example, vapor pressures for water drops and bubbles in water are given in Table 2.2.

| $r$ (nm) | $P_0^K/P_0$ drop | $P_0^K/P_0$ bubble |
|---|---|---|
| 1000 | 1.001 | 0.999 |
| 100 | 1.011 | 0.989 |
| 10 | 1.114 | 0.898 |
| 1 | 2.950 | 0.339 |

**Table 2.2:** Relative equilibrium vapor pressure of a curved water surface at $25°C$ for spherical drops and bubbles of radius $r$.

At this point it is necessary to clarify several questions which sometimes cause confusion. When do we use the term "vapor" instead of "gas"? Vapor is used when the liquid is present in the system and liquid evaporation and vapor condensation take place. This distinction is not always clear cut because, when dealing with adsorption (Chapter 9) we certainly take the two corresponding processes — adsorption and desorption — into account but still talk about gas. How does the presence of an additional background gas change the properties of a vapor? For example, does pure water vapor behave differently from water vapor at the same partial pressure in air (in the presence of nitrogen and oxygen)? Answer: To a first approximation there is no difference as long as phenomena in thermodynamic equilibrium are concerned. "First approximation" means, as long as interactions between the vapor molecules and the molecules of the background gas are negligible. However, time-dependent processes and kinetic phenomena such as diffusion can be completely different and certainly depend on the background gas. This is, for instance, the reason why drying in a vacuum is much faster than drying in air.

## 2.6 Capillary condensation

An important application of the Kelvin equation is the description of capillary condensation. This is the condensation of vapor into capillaries or fine pores even for vapor pressures below $P_0$; $P_0$ is the equilibrium vapor pressure of the liquid with a planar surface. Lord Kelvin was the one who realized that the vapor pressure of a liquid depends on the curvature of its surface. In his words this explains why "moisture is retained by vegetable substances, such as cotton cloth or oatmeal, or wheat-flour biscuits, at temperatures far above the dew point of the surrounding atmosphere" [17].

Capillary condensation can be illustrated by the model of a conical pore with a totally wetting surface (Fig. 2.12). Liquid will immediately condense in the tip of the pore. Condensation continues until the bending radius of the liquid has reached the value given by the Kelvin equation. The situation is analogous to that of a bubble and we can write

$$RT \cdot \ln \frac{P_0^K}{P_0} = -\frac{2\gamma V_m}{r_C} \tag{2.24}$$

The vapor pressure of the liquid inside the pore decreases to $P_0^K$, with $r_C$ being the capillary radius at the point where the meniscus is in equilibrium.

Many surfaces are not totally wetted, but they form a certain contact angle $\Theta$ with the liquid. In this case the radius of curvature increases. It is not longer equal to the capillary radius, but to $r = r_C/\cos\Theta$.

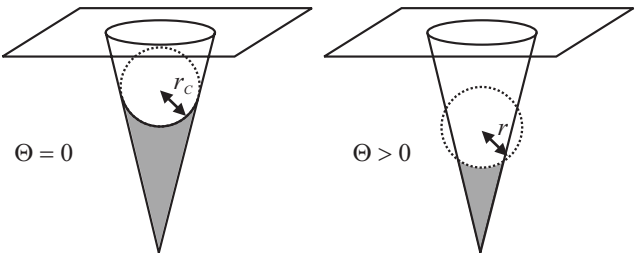



**Figure 2.12:** Capillary condensation into a conical pore with wetting and partially wetting surfaces.

■ **Example 2.6.** We have a porous solid with pores of all dimensions. It is in water vapor at 20°C. The humidity is 90%. What is the size of the pores, which fill up with water?

$$r_C = -\frac{2\gamma V_m}{RT \cdot \ln 0.9} = -\frac{2 \cdot 0.072\ \mathrm{Jm^{-2}} \cdot 18 \times 10^{-6}\mathrm{m}^3}{8.31\ \mathrm{JK^{-1}} \cdot 293\ \mathrm{K} \cdot \ln 0.9} = 10\ \mathrm{nm}$$

Attention has to be paid as to which radius is inserted into the Kelvin equation. Generally, there is no rotational symmetric geometry. Then $2/r_C$ has to be substituted by $1/R_1 + 1/R_2$. In a fissure or crack, one radius of curvature is infinitely large. Instead of $2/r_C$ there should be $1/r$ in the equation, with $r$ being the bending radius vertical to the fissure direction.

Capillary condensation has been studied by various methods, and the validity of the previous description has been confirmed for several liquids and radii of curvature down to a few nanometers [18–21].

An important consequence of capillary condensation is that liquids are strongly adsorbed into porous materials. Another important consequence is the existence of the capillary force, also called the meniscus force. Capillary condensation often strengthens the adhesion of fine particles and in many cases determines the behavior of powders. If, for instance, two particles come into contact, liquid (usually water) will condense into the gap of the contact zone [22]. The meniscus is curved. As a consequence, the Laplace pressure in the liquid is negative, and the particles attract each other.

To calculate the capillary force we take a simple but important example. Two spherical particles of identical radius $R_P$ are in contact (Fig. 2.13). We further assume that the liquid wets the surface of the particles. This is the case for clays and many other minerals, and water. The total radius of curvature of the liquid surface $1/R_1 + 1/R_2$ is

$$\frac{1}{x} - \frac{1}{r} \approx -\frac{1}{r} \tag{2.25}$$

In most practical cases we can safely assume that $x \gg r$. The pressure is therefore $\Delta P = \gamma/r$ lower in the liquid than in the outer vapor phase. It acts upon a cross-sectional area $\pi x^2$ leading

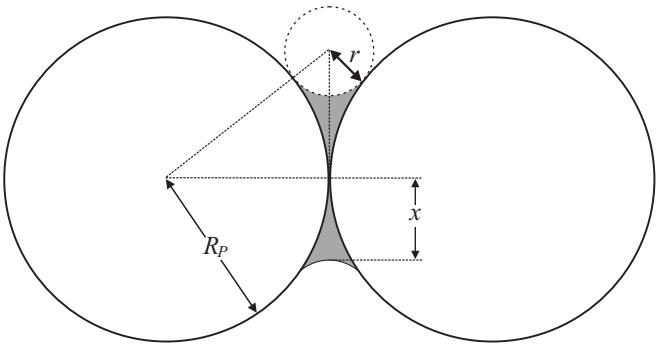



**Figure 2.13:** Two spherical particles with liquid meniscus.

to an attractive force of $\pi x^2 \Delta P$. We use Pythagoras' theorem to express $x^2$ by $r$:

$$\begin{aligned}(R_P + r)^2 = (x + r)^2 + R_P^2 \Rightarrow R_P^2 + 2rR_P + r^2 = x^2 + 2xr + r^2 + R_P^2 \\ \Rightarrow 2rR_P = x^2 + 2xr \approx x^2\end{aligned} \tag{2.26}$$

For the last approximation we assumed that $x \gg 2r$. From this follows $x^2 = 2rR_P$. Therefore the attractive capillary force is

$$F = 2\pi\gamma R_P \tag{2.27}$$

The force only depends on the radius of the particles and the surface tension of the liquid. It does not depend on the actual radius of curvature of the liquid surface nor on the vapor pressure! This is at first sight a surprising result, and is due to the fact that, with decreasing vapor pressure the radius of curvature, and therefore also $x$, decreases. At the same time the Laplace pressure increases by the same amount.

■ **Example 2.7.** A quartz sphere hangs on a second similar sphere. Some water vapor is in the room which leads to a capillary force. Small particles are held by the capillary force, large particles fall down due to the dominating gravitational force. Beyond which particle radius is gravity strong enough to separate the two spheres? For quartz ($SiO_2$) we assume a density of $\rho = 3000\ \mathrm{kgm^{-3}}$. The weight of the sphere is

$$\frac{4}{3}\pi R_P^3 \rho g = 1.23 \times 10^5 \frac{\mathrm{kg}}{\mathrm{m^2 s^2}} \cdot R_P^3$$

The capillary force is

$$2\pi \cdot 0.072\,\frac{\mathrm{J}}{\mathrm{m^2}} \cdot R_P = 0.45 \frac{\mathrm{kg}}{\mathrm{s^2}} \cdot R_P$$

Both are equal for

$$R_P = \sqrt{\frac{0.45}{1.23 \times 10^5}\,\mathrm{m^2}} = 1.9\ \mathrm{mm}$$

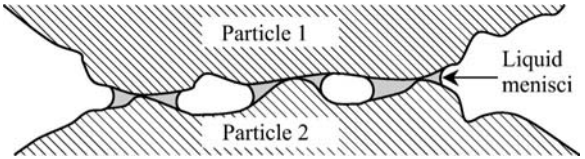



**Figure 2.14:** Two particles with rough surfaces in contact.

In reality the capillary force is often much smaller than the calculated value. This can be explained by the roughness of the surfaces. The particle surfaces are usually rough and touch only at some points. Capillary condensation takes place only at these points, as illustrated in Fig. 2.14.

The surface tension of the liquid itself gives a further direct contribution to the attraction of $F = 2\pi x\gamma$. Since $x \ll R_P$ this contribution is usually small. A calculation of the distance dependence of the capillary force and of adhesion can be found in Refs. [23–25].

## 2.7 Nucleation theory

The change of vapor pressure with the curvature of a liquid surface has important consequences for condensation and the formation of bubbles during boiling. The formation of a new phase in the absence of an external surface is called homogeneous nucleation. In homogeneous nucleation first small clusters of molecules are formed. These clusters grow due to the condensation of other molecules. In addition, they aggregate to form larger clusters. Finally macroscopic drops form. Usually this happens only if the vapor pressure is significantly above the saturation vapor pressure. In most practical situations we encounter heterogeneous nucleation, where a vapor condenses onto a surface such as a dust particle. A well known example of heterogeneous nucleation is the formation of bubbles when pouring sparkling water (or if you prefer beer) into a glass. Bubbles nucleate at the glass surface, grow in size and eventually rise.

Here we only discuss homogeneous nucleation. Though it is less common, the mathematical treatment and the concepts developed are important and are also used for other applications. The classical theory of homogeneous nucleation was developed around 1920–1935 [26, 27]. In order to describe nucleation, we calculate the change in the Gibbs free energy for the condensation of $n$ moles vapor at a vapor pressure $P$, into a drop. Please note that $n$ is much smaller than one. Keep also in mind that in this chapter, $P$ is not the total pressure. The total pressure might be higher than the vapor pressure due to the presence of other gases. To calculate the change in Gibbs free energy we first consider $G_L - G_V$. Here, $G_L$ is the Gibbs free energy of the liquid drop and $G_V$ is the Gibbs free energy of the corresponding number of molecules in the vapor phase. $G_V$ is given by

$$G_V = nG^0 + nRT \cdot \ln P \tag{2.28}$$

assuming that the vapor behaves as an ideal gas. It is more difficult to obtain an expression for $G_L$. To calculate $G_L$ we use the fact that it is equal to the Gibbs free energy of a (hypothetical) vapor, which is in equilibrium with the liquid drop. This hypothetical vapor has a pressure $P_0^K$ and its Gibbs free energy is

$$G_L = G_V^{hypothetical} = nG^0 + nRT \cdot \ln P_0^K \tag{2.29}$$

Since the drops have a curved surface of radius $r$ the vapor pressure $P_0^K$ is higher than that of the flat liquid surface. Thus, the difference in the Gibbs energies is

$$G_L - G_V = nRT \cdot \ln P_0^K - nRT \cdot \ln P = -nRT \cdot \ln \frac{P}{P_0^K} \tag{2.30}$$

This, however, is not the whole energy difference. In addition, the drop has a surface tension which has to be considered. The total change in the Gibbs free energy is

$$\Delta G = -nRT \cdot \ln \frac{P}{P_0^K} + 4\pi\gamma r^2 \tag{2.31}$$

In a drop of radius $r$ there are $n = 4\pi r^3/3V_m$ moles of molecules, where $V_m$ is the molar volume of the liquid phase. Inserting leads to

$$\Delta G = -\frac{4\pi RTr^3}{3V_m} \cdot \ln \frac{P}{P_0^K} + 4\pi\gamma r^2 \tag{2.32}$$

This is the change in Gibbs free energy upon condensation of a drop from a vapor phase with partial pressure $P$.

Let us analyse Eq. (2.32) in more detail. For $P < P_0^K$, the first term is positive and therefore $\Delta G$ is positive. Any drop, which is formed by randomly clustering molecules will evaporate again. No condensation can occur. For $P > P_0^K$, $\Delta G$ increases with increasing radius, has a maximum at the so-called critical radius $r^*$ and then decreases again. At the maximum we have $d\Delta G/dr = 0$, which leads to a critical radius of

$$r^* = \frac{2V_m\gamma}{RT \cdot \ln\left(P/P_0^K\right)} \tag{2.33}$$

One (at first sight) surprising fact, is that Eq. (2.32) is equal to the Kelvin equation (2.18). The Kelvin equation applies to systems in thermodynamic equilibrium. Since $d\Delta G/dr = 0$ the system is formally in equilibrium.

As an example, Fig. 2.15 shows a plot of $\Delta G$ versus the drop radius for water at different supersaturations. Supersaturation is the actual vapor pressure $P$ divided by the vapor pressure $P_0$ of a vapor, which is in equilibrium with a liquid having a planar surface.

■ **Example 2.8.** For water at $T = 0°$C and a supersaturation $P/P_0 = 4$, the critical radius is $R^* = 8$ Å. This corresponds approximately to 70 molecules. There $\Delta G_{max} = 1.9 \times 10^{-19}$ J.

How does nucleation proceed? In a vapor there are always a certain number of clusters. Most of them are very small and consist only of a few molecules. Others are a little larger. When the actual partial pressure $P$ becomes higher than the equilibrium vapor pressure $P_0$, large clusters occur more frequently. If a cluster exceeds the critical size, thermal fluctuations tend to enlarge it even more, until it becomes "infinitely" large and the liquid condenses.

The aim of any theory of nucleation is to find a rate $J$ with which clusters of critical size are formed. This number is proportional to the Boltzmann factor $\exp(-\Delta G_{max}/k_BT)$. A

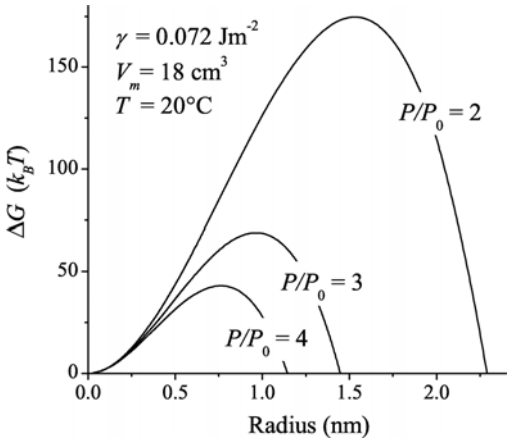



**Figure 2.15:** Change of Gibbs free energy for the condensation of vapor to a drop of a certain radius.

complete description of the classical theory of nucleation is not possible within this book. The result is:

$$J = \sqrt{\frac{2\gamma}{\pi m}} \cdot v_m \cdot \left(\frac{P}{k_B T}\right)^2 \cdot \exp\left[-\frac{16\pi v_m^2 \gamma^3}{3 \cdot (k_B T)^3 \cdot \ln^2 (P/P_0)}\right] \tag{2.34}$$

Here, $m$ is the mass and $v_m = V_m/N_A$ the volume of one molecule.

Classical nucleation theory is the basis for understanding condensation and it predicts the dependencies correctly. Unfortunately, quantitatively the predictions often do not agree with experimental results [28, 29]. Theory predicts too low nucleation rates at low temperatures. At high temperatures the calculated rates are too high. Empirical correction functions can be used and then very good agreement is achieved [30]. Ref. [31] reviews experimental methods. General overviews are Refs. [32–34].

Experimentally nucleation rates can be determined in expansion chambers [35]. The vapor is expanded in a fast and practically adiabatic process. Then it cools down. Since at low temperatures, the equilibrium vapor pressure is much lower, supersaturation is reached. Partially, this is compensated for by the pressure reduction during the expansion, but the temperature effect dominates. The density of nuclei can be measured by light scattering.

■ **Example 2.9.** The nucleation of water is analysed in an expansion chamber. A vapor at an initial pressure of 2330 Pa at 303 K is expanded to a final pressure of 1575 Pa. In this process it cools down to 260 K. At 260 K the equilibrium vapor pressure is 219 Pa. Thus, the supersaturation reaches $P/P_0 = 7.2$. What is the nucleation rate?

At 260 K the surface tension extrapolated from values above 0°C is $\gamma \approx 77$ mNm$^{-1}$. The molecular volume is $v_m = m/\rho = 2.99\times10^{-26}$ kg/1000 kgm$^{-3}$ $= 2.99\times10^{-29}$ m$^3$,

where $m$ is the mass of a water molecule. Inserting these values into Eq. (2.34) leads to a nucleation rate of

$$\begin{aligned} J &= \sqrt{\frac{2 \cdot 0.077\,\mathrm{Nm}^{-1}}{\pi \cdot 2.99 \times 10^{-26}\,\mathrm{kg}}} \cdot 2.99 \times 10^{-29}\,\mathrm{m}^3 \cdot \left(\frac{1575\,\mathrm{Pa}}{3.59 \times 10^{-21}\,\mathrm{J}}\right)^2 \\ &\quad \cdot \exp\left\{-\frac{16\pi \cdot \left(2.99 \times 10^{-29}\,\mathrm{m}^3\right)^2 \cdot \left(0.077\,\mathrm{Nm}^{-1}\right)^3}{3 \cdot \left(3.59 \times 10^{-21}\,\mathrm{J}\right)^3 \cdot \ln^2 7.2}\right\} \\ &= 1.28 \times 10^{12}\,\mathrm{s}^{-1} \cdot 2.99 \times 10^{-29}\,\mathrm{m}^3 \cdot 1.92 \times 10^{47}\,\mathrm{m}^{-6} \cdot e^{-37.9} \\ &= 2.54 \times 10^{14}\,\mathrm{s}^{-1}\mathrm{m}^{-3} \end{aligned}$$

In most practical situations nucleation occurs at certain nucleation sites [36]. One example is the formation of bubbles in champagne [37]. At the end of the fermentation process, the $CO_2$ pressure in a bottle of champagne is around 6 atm. When the bottle is opened, the pressure in the vapor phase suddenly drops and an oversaturation of typically 5 is reached. After pouring the champagne into a glass the dissolved $CO_2$ molecules escape by forming bubbles (only a small part escapes by diffusion to the surface). Several kinds of particles, which are stuck on the glass wall, are able to entrap gas pockets during the filling of the glass. These particles are responsible for the repetitive production of bubbles rising in the form of bubble trains (Fig. 2.16). Most of these particles are cellulose fibres coming form the surrounding air or remaining from the wiping process.

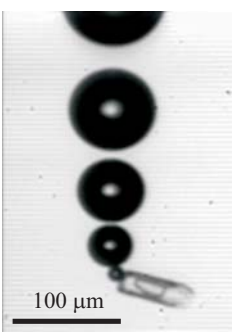



**Figure 2.16:** $CO_2$ bubbles nucleating from champagne at the bottom of a glass. Here, gas pockets entrapped inside cellulose particles serve as nucleation sites. The images, taken with a high speed video microscope, were kindly provided by G. Liger-Belair [37].

## 2.8 Summary

- The surface tension of a liquid is defined as the work required to produce a new surface per unit area:

$$dW = \gamma \cdot dA$$

- Surface tensions of liquids are typically 20–80 mN/m.
- In equilibrium and neglecting gravity, the curvature of a liquid surface is constant and given by the Young–Laplace equation:

$$\Delta P = \gamma \cdot \left(\frac{1}{R_1} + \frac{1}{R_2}\right)$$

For a liquid surface with a net curvature there is always a pressure difference across the interface. The pressure on the concave side is higher.

- Important techniques to measure the surface tension of liquids are the sessile drop method, the pendant or sessile bubble method, the Du-Noüy ring tensiometer, and the Wilhelmy-plate method.

- The vapor pressure of a liquid depends on the curvature of its surface. For drops it is increased compared to the vapor pressure of a planar surface under the same conditions. For bubbles it is reduced. Quantitatively this is described by the Kelvin equation.

- One consequence of the curvature dependence of the vapor pressure is capillary condensation, that is the spontaneous condensation of liquids into pores and capillaries. Capillary condensation plays an important role for the adsorption of liquids into porous materials and powders. It also causes the adhesion of particles. The condensing liquid forms a meniscus around the contact area of two particles which causes the meniscus force .

## 2.9 Exercises

1. We would like to study a clean solid surface. Lets assume we have produced a pure, clean surface in UHV (ultrahigh vacuum). We would like to analyze it for 1 h and we can tolerate a contamination of 10% of a monolayer. To which value do we have to reduce the pressure in the UHV chamber? Give only an estimation. You can assume that on a clean solid surface, most of the gas molecules which hit the surface are adsorbed.

2. To apply the sitting or pendant drop method to measure the surface tension of a liquid the drop must be large enough so that gravitation plays a significant role. Why?

3. A plastic box is filled with water to a height $h = 1$ m. A hole of radius 0.1 mm is drilled into the bottom. Does all water run out? The plastic is nonwetting.

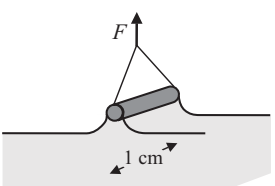
4. Wilhelmy-plate method. What is the force on a plate of 1 cm width having a contact angle of 45° in water?

5. Drop-weight method. To determine the surface tension of a hexadecane ($C_{16}H_{34}$) you let it drop out of a capillary with 4 mm outer and 40 $\mu$m inner diameter. Hexadecane wets the capillary. Its density is 773 kg/m$^3$. 100 drops weigh 2.2 g. Calculate the surface tension of hexadecane using the simple Eq. (2.15) and the correction factor $f$. It was concluded that $f$ should be a function of $r_C/V^{1/3}$, with $V$ being the volume of the drop. Values for the correction factor are listed in the following table (from Ref. [1], p. 19). Is it necessary to use the correction?

| $r_c/V^{1/3}$ | $f$ | $r_c/V^{1/3}$ | $f$ | $r_c/V^{1/3}$ | $f$ |
|---|---|---|---|---|---|
| 0.00 | 1.0000 | 0.75 | 0.6032 | 1.25 | 0.652 |
| 0.30 | 0.7256 | 0.80 | 0.6000 | 1.30 | 0.640 |
| 0.35 | 0.7011 | 0.85 | 0.5992 | 1.35 | 0.623 |
| 0.40 | 0.6828 | 0.90 | 0.5998 | 1.40 | 0.603 |
| 0.45 | 0.6669 | 0.95 | 0.6034 | 1.45 | 0.583 |
| 0.50 | 0.6515 | 1.00 | 0.6098 | 1.50 | 0.567 |
| 0.55 | 0.6362 | 1.05 | 0.6179 | 1.55 | 0.551 |
| 0.60 | 0.6250 | 1.10 | 0.6280 | 1.60 | 0.535 |
| 0.65 | 0.6171 | 1.15 | 0.6407 | | |
| 0.70 | 0.6093 | 1.20 | 0.6535 | | |

6. A hydrophilic sphere of radius $R_P = 5\,\mu$m sits on a hydrophilic planar surface. Water from the surrounding atmosphere condenses into the gap. What is the circumference of the meniscus? Make a plot of radius of circumference $x$ versus humidity. At equilibrium the humidity is equal to $P_0^K/P_0$.

7. To measure the surface tension of a liquid a wire of $r = 1$ mm radius and $l = 1$ cm length is dipped into the liquid parallel to the liquid surface. The force required to pull the wire out of the completely wetting liquid is measured to be 0.49 mN. What is the surface tension of the liquid? What if the liquid forms a certain contact angle $\Theta > 90°$ with the solid surface of the wire? Derive the relationship between force, surface tension, and contact angle.



# 3 Thermodynamics of interfaces

In this chapter we introduce the basic thermodynamics of interfaces. The purpose is to present some important equations, learn to apply them, provide a broader base of understanding, and point out some of the difficulties. For a thorough understanding, further reading is certainly necessary. A good introduction into the thermodynamics of interfaces is given in the book of Lyklema [7].

## 3.1 The surface excess

The presence of an interface influences generally all thermodynamic parameters of a system. To consider the thermodynamics of a system with an interface, we divide that system into three parts: The two bulk phases with volumes $V^\alpha$ and $V^\beta$, and the interface $\sigma$.

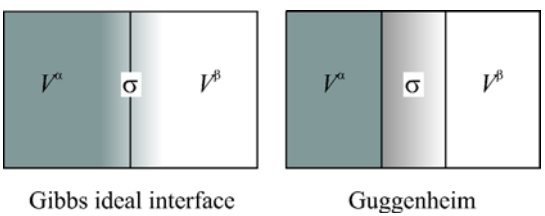



**Figure 3.1:** Left: In Gibbs convention the two phases $\alpha$ and $\beta$ are separated by an ideal interface $\sigma$ which is infinitely thin. Right: Guggenheim explicitly treated an extended interphase with a volume.

In this introduction we adhere to the Gibbs[1] convention. In this convention the two phases $\alpha$ and $\beta$ are thought to be separated by an infinitesimal thin boundary layer, the **Gibbs dividing plane**. This is of course an idealization and the Gibbs dividing plane is also called an ideal interface. There are alternative models. Guggenheim, for example, takes the extended interfacial region, including its volume, explicitly into account [38, 39]. We use the Gibbs model because in most applications it is more practical.

In the Gibbs model the interface is ideally thin ($V^\sigma = 0$) and the total volume is

$$V = V^\alpha + V^\beta \tag{3.1}$$

[1] Josiah Willard Gibbs, 1839–1903. American mathematician and physicist, Yale College.

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

All other extensive quantities can be written as a sum of three components: one of bulk phase $\alpha$, one of bulk phase $\beta$, and one of the interfacial region $\sigma$. Examples are the internal energy $U$, the number of molecules of the $i$th substance $N_i$, and the entropy $S$:

$$U = U^\alpha + U^\beta + U^\sigma \tag{3.2}$$

$$N_i = N_i^\alpha + N_i^\beta + N_i^\sigma \tag{3.3}$$

$$S = S^\alpha + S^\beta + S^\sigma \tag{3.4}$$

The contributions of the two phases and of the interface are derived as follows. Let $u^\alpha$ and $u^\beta$ be the internal energies per unit volume of the two phases. The internal energies $u^\alpha$ and $u^\beta$ are determined from the homogeneous bulk regions of the two phases. Close to the interface they might be different. Still, we take the contribution of the volume phases to the total energy of the system as $u^\alpha V^\alpha + u^\beta V^\beta$. The internal energy of the interface is

$$U^\sigma = U - u^\alpha V^\alpha - u^\beta V^\beta \tag{3.5}$$

At an interface, the molecular constitution changes. The concentration (number of molecules per unit volume) of the $i^{th}$ material is, in the two phases, respectively $c_i^\alpha$ and $c_i^\beta$. The additional quantity that is present in the system due to the interface is

$$N_i^\sigma = N_i - c_i^\alpha V^\alpha - c_i^\beta V^\beta \tag{3.6}$$

With Eq. (3.6) it is possible to define something like a surface concentration, the so called **interfacial excess**:

$$\Gamma_i = \frac{N_i^\sigma}{A} \tag{3.7}$$

$A$ is the interfacial area. The interfacial excess is given as a number of molecules per unit area ($\mathrm{m}^{-2}$) or in mol/m$^2$.

In the Gibbs model of an ideal interface there is one problem: where precisely do we position the ideal interface? Let us therefore look at a liquid–vapor interface of a pure liquid more closely. The density decreases continuously from the high density of the bulk liquid to the low density of the bulk vapor (see Fig. 3.2). There could even be a density maximum in between since it should in principle be possible to have an increased density at the interface. It is natural to place the ideal interface in the middle of the interfacial region so that $\Gamma = 0$. In this case the two dotted regions, left and right from the ideal interface, are equal in size. If the ideal interface is placed more into the vapor phase the total number of molecules extrapolated from the bulk densities is higher than the real number of molecules, $N < c^\alpha V^\alpha + c^\beta V^\beta$. Therefore the surface excess is negative. Vice versa: if the ideal interface is placed more into the liquid phase, the total number of molecules extrapolated from the bulk densities is lower than the real number of molecules, $N > c^\alpha V^\alpha + c^\beta V^\beta$, and the surface excess is positive.

Let us now turn to two- or multi-component liquids such as a solvent with dissolved substances. Substituting $V^\alpha = V - V^\beta$ we can write

$$N_1^\sigma = N_1 - c_1^\alpha V + \left(c_1^\alpha - c_1^\beta\right) V^\beta \tag{3.8}$$

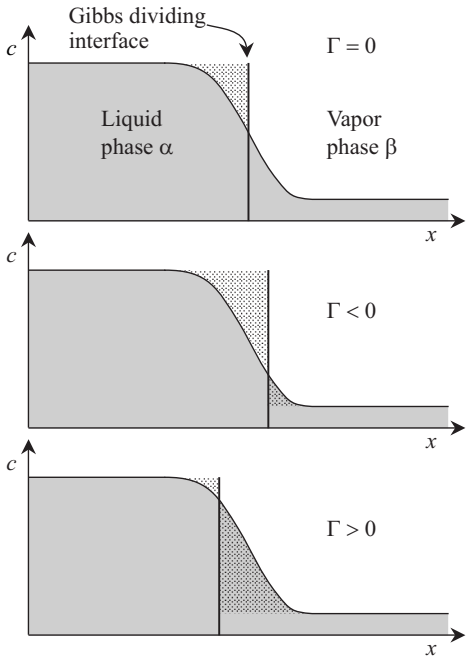



**Figure 3.2:** Dependence of the surface excess Γ on the position of the Gibbs dividing plane.

for the first component which is taken to be the solvent. For all other components we get similar equations.

$$N_i^\sigma = N_i - c_i^\alpha V + \left(c_i^\alpha - c_i^\beta\right) V^\beta \tag{3.9}$$

All quantities on the right side of the equations, except $V^\beta$, do not depend on the position of the dividing plane and are measurable quantities. Only $V^\beta$, depends on the choice of the dividing plane. We can eliminate $V^\beta$ by multiplying Eq. (3.8) by $(c_i^\alpha - c_i^\beta)/(c_1^\alpha - c_1^\beta)$ and subtracting Eq. (3.8) form Eq. (3.9):

$$N_i^\sigma - N_1^\sigma \frac{c_i^\alpha - c_i^\beta}{c_1^\alpha - c_1^\beta} = N_i - c_i^\alpha V - (N_1 - c_1^\alpha V) \frac{c_i^\alpha - c_i^\beta}{c_1^\alpha - c_1^\beta} \tag{3.10}$$

The right side of the equation does not depend on the position of the Gibbs dividing plane and thus, also, the left side is invariant. We divide this quantity by the surface area and obtain the invariant quantity

$$\Gamma_i^{(1)} \equiv \Gamma_i^\sigma - \Gamma_1^\sigma \frac{c_i^\alpha - c_i^\beta}{c_1^\alpha - c_1^\beta} \tag{3.11}$$

It is called **relative adsorption** of component $i$ with respect to component 1. This is an important quantity because it can be determined experimentally! As we shall see later it can be measured by determining the surface tension of a liquid versus the concentration of the solute.

■ **Example 3.1.** To show how our choice of the position of the Gibbs dividing plane influences the surface excess , we consider an equimolar mixture of ethanol and water (p. 25 of Ref. [40]). If the position of the ideal interface is such that $\Gamma_{H_2O} = 0$, one finds experimentally that $\Gamma_{Ethanol} = 9.5 \times 10^{-7}$ mol/m$^2$. If the interface is placed 1 nm outward, then we obtain $\Gamma_{Ethanol} = -130 \times 10^{-7}$ mol/m$^2$.

For the case when component 1 is a solvent in which all other components are dissolved and thus have a much lower concentration than component 1, we choose the position of the dividing plane such that $\Gamma_1^\sigma = 0$ and from Eq. (3.11) we get

$$\Gamma_i^{(1)} = \Gamma_i^\sigma \tag{3.12}$$

In Fig. 3.3 the concentration profiles for solute 2 dissolved in liquid 1 are illustrated. We assume that the solute is enriched at the surface. The area of the dotted region corresponds to the surface excess $\Gamma_2^{(1)}$ of solute.

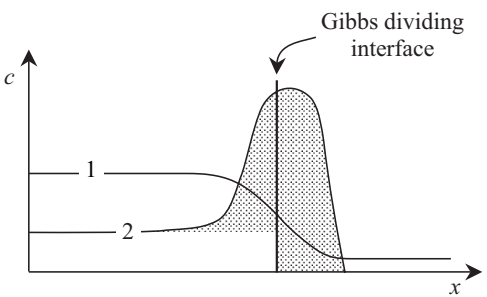



**Figure 3.3:** Concentration profile of a solute (2) dissolved in a liquid (1). The area of the dotted region corresponds to the surface excess $\Gamma_2^{(1)}$ of solute.

## 3.2 Fundamental thermodynamic relations

### 3.2.1 Internal energy and Helmholtz energy

Let us consider a process in a system with two phases, $\alpha$ and $\beta$, which are divided by an interface; we could, for instance, do work on that system. As a consequence the state quantities like the internal energy, the entropy, etc. change. How do they change and how can we describe this mathematically? In contrast to the usual "bulk" thermodynamics we have to take the interface into account.

We start the analysis with the internal energy. A variation of the internal energy of a two-phase system is, according to the first and second principle of thermodynamics,

$$dU = TdS - PdV + \sum \mu_i dN_i + dW \tag{3.13}$$

Here, $W$ is the work done on the system without expansion work $PdV$. It contains the surface work $\gamma dA$. The sum runs over all components, that means over all substances that are chemically different. $\mu_i$ is the chemical potential of the $i$th substance.

We first analyze the internal energy, and not the enthalpy, the free energy, or the Gibbs free energy, because the internal energy only contains extensive quantities $(S, V, N_i, A)$ as

variables. This simplifies the following calculation. We split the internal energy:

$$\begin{aligned} dU &= dU^\alpha + dU^\beta + dU^\sigma \\ &= TdS^\alpha + \sum \mu_i^\alpha dN_i^\alpha - P^\alpha dV^\alpha + TdS^\beta \\ &+ \sum \mu_i^\beta dN_i^\beta - P^\beta dV^\beta + TdS^\sigma + \sum \mu_i^\sigma dN_i^\sigma + \gamma dA \end{aligned} \tag{3.14}$$

The $TdS$ terms stands for the change in internal energy, which is caused by an entropy change, e.g. a heat flow. The $\mu_i dN_i$ terms consider the energy change caused by a change in the composition. Both $PdV$ terms correspond to the volume-work of the two phases. Since the interface is infinitely thin it cannot perform volume work.

With $dV = dV^\alpha + dV^\beta \Rightarrow dV^\alpha = dV - dV^\beta$ and summing up the entropy terms, the equation simplifies as:

$$\begin{aligned} dU &= TdS - P^\alpha dV - \left(P^\beta - P^\alpha\right) dV^\beta \\ &+ \sum \mu_i^\alpha dN_i^\alpha + \sum \mu_i^\beta dN_i^\beta + \sum \mu_i^\sigma dN_i^\sigma + \gamma dA \end{aligned} \tag{3.15}$$

Now we consider the Helmholtz free energy. The change in free energy of the system is $dF = -SdT - PdV + \sum \mu_i dN_i + dW$. It follows that

$$\begin{aligned} dF &= -SdT - P^\alpha dV - \left(P^\beta - P^\alpha\right) dV^\beta \\ &+ \sum \mu_i^\alpha dN_i^\alpha + \sum \mu_i^\beta dN_i^\beta + \sum \mu_i^\sigma dN_i^\sigma + \gamma dA \end{aligned} \tag{3.16}$$

When the temperature and volume are constant ($dV = 0$, $dT = 0$) the first two terms are zero.

### 3.2.2 Equilibrium conditions

In equilibrium Eq. 3.16 can be simplified even further because the chemical potentials in the three phases are equal. This we can easily demonstrate. Therefore we assume that there is no exchange of material with the outside world ($dN_i = 0$); we have a closed system. Then the three parameters $N_i^\alpha$, $N_i^\beta$, and $N_i^\sigma$ are not independent because $N_i = N_i^\alpha + N_i^\beta + N_i^\sigma$ is constant. Only two at a time, as an example $N_i^\alpha$ and $N_i^\beta$, can be varied independently. $N_i^\sigma$ is then determined by the other two amounts because $dN_i^\sigma = -dN_i^\alpha - dN_i^\beta$. Therefore we can write:

$$dF = -\left(P^\beta - P^\alpha\right) dV^\beta + \gamma dA + \sum (\mu_i^\alpha - \mu_i^\sigma) dN_i^\alpha + \sum (\mu_i^\beta - \mu_i^\sigma) dN_i^\beta \tag{3.17}$$

At equilibrium, with constant volume, temperature, and constant amounts of material, the free energy is minimal. At a minimum the derivatives with respect to all independent variables must be zero:

$$\frac{dF}{dN_i^\alpha} = \mu_i^\alpha - \mu_i^\sigma = 0, \quad \frac{dF}{dN_i^\beta} = \mu_i^\beta - \mu_i^\sigma = 0 \tag{3.18}$$

It follows that

$$\mu_i^\alpha = \mu_i^\sigma = \mu_i^\beta = \mu_i \tag{3.19}$$

Hence, in equilibrium the chemical potentials are the same everywhere in the system. With this, we can further simplify Eq. (3.17):

$$dF = -\left(P^\beta - P^\alpha\right) dV^\beta + \gamma dA + \sum \mu_i dN_i \tag{3.20}$$

This equation allows us to define the surface tension based on thermodynamics:

$$\left.\frac{\partial F}{\partial A}\right|_{T,V,V^\beta,N_i} \equiv \gamma \tag{3.21}$$

The surface tension tells us how the Helmholtz free energy of the system changes when increasing the surface area while keeping the temperature, the total volume, the volume of phase $\beta$ and the total numbers of all components constant.

Is this a useful equation? It is not so difficult to control $T$, $V$, and $N_i$ but $V^\beta$ might be difficult to keep constant. As we shall see later, for planar surfaces (and practically those which have small curvatures) the condition that $V^\beta$ has to be kept constant can be dropped.

Question: Why is it not possible, using the same argument, e.g. $dF/dA = 0$, to conclude that at equilibrium $\gamma$ must be zero? Explanation: The surface area $A$ is not an independent parameter. Surface $A$ and the volume $V^\beta$ are related. If the volume of a body changes, in general its surface area also changes. $V^\beta$ and $A$ can thus not be varied independently. In fact, a law of differential geometry says that in general $\partial V/\partial A = (1/R_1 + 1/R_2)^{-1}$.

At this point we mention a simple, alternative way of deriving the Laplace equation. In equilibrium we have $dF/dA = 0$. It leads to

$$\frac{dF}{dA} = \frac{\partial F}{\partial A} + \frac{\partial F}{\partial V^\beta} \cdot \frac{\partial V^\beta}{\partial A} = \gamma - \left(P^\beta - P^\alpha\right) \cdot \frac{\partial V^\beta}{\partial A} = 0 \tag{3.22}$$

Inserting $\partial V^\beta/\partial A = (1/R_1 + 1/R_2)^{-1}$, taking $\Delta P = P^\beta - P^\alpha$, and rearranging the equation directly leads to the Young–Laplace equation.

### 3.2.3 Location of the interface

At this point we should note that, fixing the bending radii, we define the location of the interface. A possible choice for the ideal interface is the one that is defined by the Laplace equation. If the choice for the interface is different, the value for the surface tension must be changed accordingly. Otherwise the Laplace equation would no longer be valid. All this can

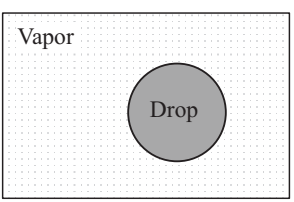



**Figure 3.4:** A drop in its vapor phase.

be illustrated with the example of a spherical drop [41]. We can, for instance, consider the evaporation or the condensation of liquid from, or to, a drop of radius $r$. There we have

$$V^\beta = \frac{4\pi}{3} r^3, \quad A = 4\pi r^2 \quad \Rightarrow V^\beta = \frac{4\pi}{3}\left(\frac{A}{4\pi}\right)^{3/2} \Rightarrow \frac{\partial V^\beta}{\partial A} = \frac{r}{2} \tag{3.23}$$

If the interface is chosen to be at a radius $r'$, then the corresponding value for $\partial V^\beta/\partial A$ is $r'/2$. The pressure difference $P^\beta - P^\alpha$ can in principle be measured. This implies that $P^\beta - P^\alpha = 2\gamma/r$ and $P^\beta - P^\alpha = 2\gamma'/r'$ are both valid at the same time. This is only possible if, dependent on the radius, one accepts a different interfacial tension. Therefore we used $\gamma'$ in the second equation. In the case of a curved surface, the interfacial tension depends on the location of the Gibbs dividing plane! In the case of flat surfaces this problem does not occur. There, the pressure difference is zero and the surface tension is independent of the location of the ideal interface.

A possible objection could be that the surface tension is measurable and thus the Laplace equation assigns the location of the ideal interface. But this is not true. The only quantity that can be measured is mechanical work and the forces acting during the process. For curved surfaces it is not possible to divide volume and surface work. Therefore, it is not possible to measure *only* the surface tension.

### 3.2.4 Gibbs energy and definition of the surface tension

In this chapter we introduce a more useful equation for the surface tension. This we do in two steps. First, we seek an equation for the change in the Gibbs free energy. The Gibbs free energy $G$ is usually more important than $F$ because its natural variables, $T$ and $P$, are constant in most applications. Second, we have just learned that, for curved surfaces, the surface tension is not uniquely defined and depends on where precisely we choose to position the interface. Therefore we concentrate on planar surfaces from now on.

For the Gibbs free energy we write

$$dG = -SdT + V^\alpha dP^\alpha + V^\beta dP^\beta + \sum \mu_i dN_i + \gamma dA \tag{3.24}$$

Assuming that the interface is flat (planar) we have the same pressure in both phases $P^\alpha = P^\beta = P$ and we get

$$dG = -SdT + VdP + \sum \mu_i dN_i + \gamma dA \tag{3.25}$$

With the help of this equation it is also possible to give a definition of the interfacial tension, which is equivalent to the previous definition:

$$\left.\frac{\partial G}{\partial A}\right|_{T,P,N_i} \equiv \gamma \tag{3.26}$$

The surface tension is the increase in the Gibbs free energy per increase in surface area at constant $T$, $P$, and $N_i$.

### 3.2.5 Free surface energy, interfacial enthalpy and Gibbs surface energy

Until now we have considered the total energy quantities of the system. Now we turn to the interfacial excess quantities. We start with the internal interfacial or internal surface energy

$$dU^\sigma = TdS^\sigma + \sum \mu_i dN_i^\sigma + \gamma dA \tag{3.27}$$

The term $PdV^\sigma$ disappears, because the ideal interface has no volume. We integrate the expression keeping the intrinsic parameters $T, \mu_i$, and $\gamma$ constant.[2] This integration is allowed because it represents, in principle, a feasible process, e.g., through simply increasing the surface area of the system. This can be realized by, for instance, tilting a sealed test tube which is partially filled with a liquid. Result:

$$U^\sigma = TS^\sigma + \sum \mu_i N_i^\sigma + \gamma A \tag{3.28}$$

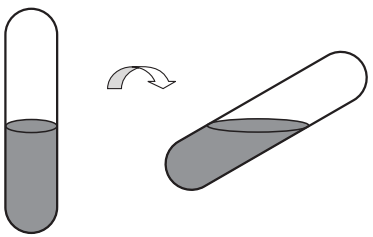


**Figure 3.5:** Increasing the surface area size by tilting a test tube.

For the free energy of the surface we use $F^\sigma = U^\sigma - TS^\sigma$ and get

$$F^\sigma = \gamma A + \sum \mu_i N_i^\sigma \Rightarrow \frac{F^\sigma}{A} = \gamma + \sum \mu_i \Gamma_i \tag{3.29}$$

The differential of $F^\sigma$ is:

$$dF^\sigma = -S^\sigma dT + \sum \mu_i dN_i^\sigma + \gamma dA \tag{3.30}$$

Before turning to the surface enthalpy we would like to derive an important relationship between the surface entropy and the temperature dependence of the surface tension. The Helmholtz interfacial free energy is a state function. Therefore we can use the Maxwell relations and obtain directly an important equation for the surface entropy:

$$-\left.\frac{\partial S^\sigma}{\partial A}\right|_{T,N_i^\sigma} = \left.\frac{\partial \gamma}{\partial T}\right|_{A,N_i^\sigma} \tag{3.31}$$

There are two common and widely used definitions of the interfacial excess enthalpy. We can argue that enthalpy is equal to the internal energy minus the total mechanical work $\gamma A - PV^\sigma$. Since in the Gibbs convention $PV^\sigma = 0$ we define

$$H^\sigma \equiv U^\sigma - \gamma A \tag{3.32}$$

This definition is recommended by the IUPAC [42]. One consequence is that $H = H^\alpha + H^\beta + H^\sigma + \gamma A$. The differential is again easily obtained to be

$$dH^\sigma = TdS^\sigma + \sum \mu_i dN_i^\sigma - Ad\gamma \tag{3.33}$$

[2] In the Gibbs model all volume terms disappear. In the Guggenheim model one must also keep $P$ constant during the integration.

Alternatively, one could argue that the enthalpy is equal to the internal energy minus the volume work $PV^\sigma$. Since the volume work is zero in the Gibbs convention we simply get

$$H'^\sigma \equiv U^\sigma \tag{3.34}$$

What is the interfacial excess Gibbs energy? The difference between $U^\sigma$ and $F^\sigma$ should be the same as the one between $H^\sigma$ and $G^\sigma$. Therefore we define

$$G^\sigma \equiv H^\sigma - TS^\sigma = F^\sigma - \gamma A = \sum \mu_i N_i^\sigma \tag{3.35}$$

One consequence is that $G = G^\alpha + G^\beta + G^\sigma + \gamma A$. The differential is

$$dG^\sigma = -S^\sigma dT + \sum \mu_i dN_i^\sigma - A d\gamma \tag{3.36}$$

With the alternative definition of $H'^\sigma$ we obtain

$$G'^\sigma \equiv H'^\sigma - TS^\sigma = F^\sigma \tag{3.37}$$

and $G = G^\alpha + G^\beta + G'^\sigma$.

## 3.3 The surface tension of pure liquids

For pure liquids the description becomes much simpler. We start by asking, how is the surface tension related to the surface excess quantities, in particular to the internal surface energy and the surface entropy?

One important relationship can be derived directly from Eq. (3.29). For pure liquids we choose the Gibbs dividing plane such that $\Gamma = 0$. Then the surface tension is equal to the free surface energy per unit area:

$$f^\sigma = \frac{F^\sigma}{A} = \gamma \tag{3.38}$$

Let us turn to the entropy. We start with Eq. (3.31). For pure liquids the position of the interface is chosen such that $N^\sigma = 0$. For homogeneous systems we also know that $s^\sigma \equiv S^\sigma/A = \partial S^\sigma/\partial A$. Putting everything together we find

$$s^\sigma = -\left.\frac{\partial \gamma}{\partial T}\right|_{P,A} \tag{3.39}$$

The surface entropy per unit area is given by the change in the surface tension with temperature. In order to determine the surface entropy one needs to measure how the surface tension changes with temperature.

Question: If the volume of the interface is zero, why is the condition important that $P$ is constant? Reason: A change in pressure might change the quality of the interface and thus its entropy. This equation is generally valid, not only within the Gibbs formalism.

For the majority of liquids, the surface tension decreases with increasing temperature. This behaviour was already observed by Eötvös, Ramsay & Shields at the end of the last century [43,44]. The entropy on the surface is thus increased, which implies that the molecules at the surface are less ordered than in the bulk liquid phase.

What about the internal energy? For a pure liquid we have $U^\sigma = TS^\sigma + \gamma A$. Division by $A$ and, assuming that we have a homogeneous system, leads to

$$u^\sigma \equiv \frac{U^\sigma}{A} = Ts^\sigma + \gamma \tag{3.40}$$

or

$$u^\sigma = \gamma - T \cdot \left.\frac{\partial \gamma}{\partial T}\right|_{P,A} \tag{3.41}$$

It is thus possible to determine the internal surface energy and the surface entropy by measuring the dependence of the surface tension on the temperature.

■ **Example 3.2.** The surface tension of water decreases with increasing temperature from 74.23 mN/m at 10°C, to 71.99 at 25°C, and 67.94 mN/m at 50°C. Calculate $f^\sigma$, $s^\sigma$, and $u^\sigma$ at 25°C.

The first one is easy to answer: $f^\sigma = \gamma = 71.99$ mNm$^{-1}$. Using the other two values for the surface tension we can estimate that $\partial\gamma/\partial T = 157.3 \times 10^{-6}$ Nm$^{-1}$K$^{-1}$ at 25°C. Thus, the surface entropy is $s^\sigma = -157.3 \times 10^{-6}$ Nm$^{-1}$K$^{-1}$. Using Eq. (3.41) the internal surface energy is obtained to be $u^\sigma = 71.99 + 293.2 \cdot 0.1573$ Nm$^{-1}$ $= 118.09$ nm$^{-1}$.

How does the heat flow during an increase in the surface area? In a reversible process $TdS$ is the heat $\delta Q$ that the system absorbs. The heat absorption is proportional to the surface increase and we can write $\delta Q = qdA$. Here, $q$ is the heat per unit area that is taken up by the system. With $dS = s^\sigma dA$ and $s^\sigma = -\partial\gamma/\partial T$ we get

$$qdA = \delta Q = TdS = Ts^\sigma dA = -T \cdot \frac{\partial \gamma}{\partial T} \cdot dA \tag{3.42}$$

or

$$q = -T \cdot \frac{\partial \gamma}{\partial T} \tag{3.43}$$

This is the heat per unit area absorbed by the system during an isothermal increase in the surface. Since $\partial\gamma/\partial T$ is mostly negative the system usually takes up heat when the surface area is increased. Table 3.1 lists the surface tension, surface entropy, surface enthalpy, and internal surface energy of some liquids at 25°C.

## 3.4 Gibbs adsorption isotherm

It is well known that the surface tension of water decreases when a detergent is added. Detergents are strongly enriched at the surface, which lowers the surface tension. This change of surface tension upon adsorption of substances to the interface, is described by the Gibbs adsorption isotherm.

**Table 3.1:** Surface tension, surface entropy, surface enthalpy, and internal surface energy of some liquids at 25°C.

| | $\gamma = f^\sigma(\mathrm{mNm}^{-1})$ | $T \cdot s^\sigma(\mathrm{mNm}^{-1})$ | $u^\sigma(\mathrm{mNm}^{-1})$ |
|---|---|---|---|
| Mercury | 485.48 | 61.1 | 549.6 |
| Water | 71.99 | 46.9 | 121.1 |
| $n$-hexane | 17.89 | 30.5 | 49.9 |
| $n$-heptane | 19.65 | 29.2 | 50.3 |
| $n$-octane | 21.14 | 28.3 | 50.9 |
| $n$-nonane | 22.38 | 27.9 | 51.7 |
| $n$-decane | 23.37 | 27.4 | 52.2 |
| Methanol | 22.07 | 23.0 | 46.3 |
| Ethanol | 21.97 | 24.8 | 48.0 |
| 1-propanol | 23.32 | 23.1 | 47.6 |
| 1-butanol | 24.93 | 26.8 | 53.0 |
| 1-hexanol | 23.81 | 29.8 | 55.6 |
| Toluene | 27.93 | 35.4 | 65.1 |

### 3.4.1 Derivation

The Gibbs adsorption isotherm is a relationship between the surface tension and the excess interfacial concentrations. To derive it we start with Eqs. (3.27) and (3.28). Differentiation of Eq. (3.28) leads to

$$dU^\sigma = TdS^\sigma + S^\sigma dT + \sum \mu_i dN_i^\sigma + \sum N_i^\sigma d\mu_i + \gamma dA + Ad\gamma \tag{3.44}$$

Equating this to expression (3.27) results in

$$0 = S^\sigma dT + \sum N_i^\sigma d\mu_i + Ad\gamma \tag{3.45}$$

At constant temperature it can be simplified to

$$d\gamma = -\sum \Gamma_i d\mu_i \tag{3.46}$$

Equations (3.45) and (3.46) are called **Gibbs adsorption isotherms**. In general, "isotherms" are state functions plotted versus pressure, concentration, etc. at constant temperature.

One word of caution: The given equation is only valid for those surfaces whose deformation is reversible and plastic, i.e., liquid surfaces. In solids, changes of the surface are usually accompanied by elastic processes [45–47]. In order to consider elastic tensions an additional term, $(\gamma_{plas} - \gamma)dA$, has to be added to Eq. (3.45):

$$0 = S^\sigma dT + \sum N_i^\sigma d\mu_i + Ad\gamma + (\gamma_{plas} - \gamma)dA \tag{3.47}$$

Here, $\gamma_{plas}$ is a proportionality factor in front of $dA$. It describes the interfacial work in the case of pure plastic, reversible work. If the actual proportionality factor $\gamma$ is different, then the last term becomes effective.

### 3.4.2 System of two components

The simplest application of the Gibbs adsorption isotherm is a system of two components, e.g., a solvent 1 and a solute 2. In this case we have

$$d\gamma = -\Gamma_1 d\mu_1 - \Gamma_2 d\mu_2 \tag{3.48}$$

The ideal interface is conveniently defined such that $\Gamma_1 = 0$. Then we get

$$d\gamma = -\Gamma_2^{(1)} d\mu_2 \tag{3.49}$$

The superscript "(1)" should remind us of the special choice of the interface. The chemical potential of the solute is described by the equation

$$\mu_2 = \mu_2^0 + RT \cdot \ln \frac{a}{a_0} \tag{3.50}$$

Here, $a$ is the activity and $a_0$ is a standard activity (1 mol/L). Differentiating with respect to $a/a_0$ at constant temperature leads to

$$d\mu_2 = RT \cdot \frac{d\,(a/a_0)}{a/a_0} = RT \cdot \frac{da}{a} \tag{3.51}$$

Substituting this into Eq. (3.49) leads to

$$\Gamma_2^{(1)} = -\frac{a}{RT} \cdot \left.\frac{\partial \gamma}{\partial a}\right|_T \tag{3.52}$$

This is a very important equation. It directly tells us that when a solute is enriched at the interface ($\Gamma_2^{(1)} > 0$), the surface tension decreases when the solution concentration is increased. Such solutes are said to be surface active and they are called **surfactants** or surface active agents. Often the term amphiphilic molecule or simply amphiphile is used. An amphiphilic molecule consist of two well-defined regions: One which is oil-soluble (lyophilic or hydrophobic) and one which is water-soluble (hydrophilic).

When a solute avoids the interface ($\Gamma_2^{(1)} < 0$), the surface tension increases by adding the substance. Experimentally Equation (3.52) can be used to determine the surface excess by measuring the surface tension versus the bulk concentration. If a decrease in the surface tension is observed, the solute is enriched in the interface. If the surface tension increases upon addition of solute, then the solute is depleted in the interface.

■ **Example 3.3.** You add 0.5 mM SDS (sodium dodecylsulfate, $NaSO_4(CH_2)_{11}CH_3$) to pure water at 25°C. This leads to a decrease in the surface tension from 71.99mJ/m$^2$ to 69.09 mJ/m$^2$. What is the surface excess of SDS?

At such low activities and as an approximation we replace the activity $a$ by the concentration $c$ and get

$$\frac{\partial \gamma}{\partial a} \approx \frac{\Delta \gamma}{\Delta c} = \frac{(0.06909 - 0.07199)\mathrm{Nm}^{-1}}{(0.0005 - 0)\mathrm{molm}^{-3}} = -5.80\ \frac{\mathrm{Nm}^2}{\mathrm{mol}} \tag{3.53}$$

It follows that

$$\Gamma = -\frac{a}{RT} \cdot \frac{\partial \gamma}{\partial a} = -\frac{0.0005\ \mathrm{molm^{-3}}}{8.31 \cdot 298\ \mathrm{Jmol^{-1}}} \cdot (-5.80)\frac{\mathrm{Nm^2}}{\mathrm{mol}} = 1.17 \times 10^{-6}\frac{\mathrm{mol}}{\mathrm{m^2}} \quad (3.54)$$

Every molecule occupies an average surface area of 1.42 $\mathrm{nm}^2$.

The choice of the ideal interface in the Gibbs adsorption isotherm (3.52) for a two-component system is, in a certain view, arbitrary. It is, however, convenient. There are two reasons: First, on the right side there are physically measurable quantities $(a, \gamma, T)$, which are related in a simple way to the interfacial excess. Any other choice of the interface would lead to a more complicated expression. Second, the choice of the interface is intuitively evident, at least for $c_1 \gg c_2$. One should, however, keep in mind that different spatial distributions of the solute can lead to the same $\Gamma_2^{(1)}$. Figure 3.6 shows two examples of the same interfacial excess concentration $\Gamma_2^{(1)}$. In the first case the distribution of molecules 2 stretches out beyond the interface, but the concentration is nowhere increased. In the second case, the concentration of the molecules 2 is actually increased.

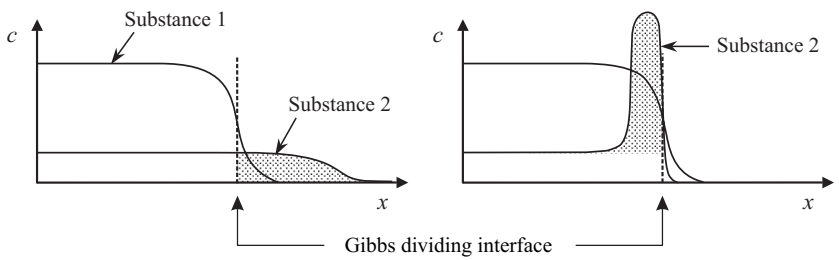



**Figure 3.6:** Examples of two different concentration profiles leading to the same interfacial excess concentration $\Gamma_2^{(1)}$.

### 3.4.3 Experimental aspects

How can Eq. (3.52) be verified? For verification, the two variables — concentration and surface tension — need to be determined independently. One way is to use radioactively labeled dissolved substances. The radioactivity close to the surface is measured. $\beta$-emitters ($^3$H, $^{14}$C, $^{35}$S) are suitable because electrons only travel a short range, i.e., any recorded radioactivity comes from molecules from the interface, or close below [48].

Plots of surface tension versus concentration for $n$-pentanol [49], LiCl (based on Ref. [50]), and SDS in an aqueous medium at room temperature are shown in Fig. 3.7. The three curves are typical for three different types of adsorption. The SDS adsorption isotherm is typical for amphiphilic substances. In many cases, above a certain critical concentration defined aggregates called micelles are formed (see Section 12.1). This concentration is called the critical micellar concentration (CMC). In the case of SDS at 25°C this is at 8.9 mM. Above the CMC the surface tension does not change significantly any further because any added substance goes into micelles not to the liquid–gas interface.

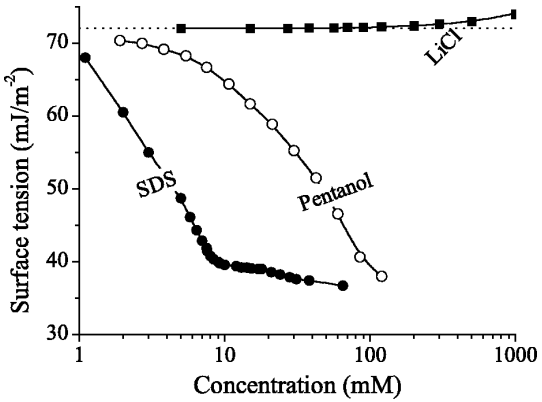



**Figure 3.7:** Plots of surface tension versus concentration for $n$-pentanol [49], LiCl (based on ref. [50]), and SDS in an aqueous medium at room temperature.

The adsorption isotherm for pentanol is typical for lyophobic substances, i.e., substances which do not like to stay in solution, and for weakly amphiphilic substances. They become enriched in the interface and decrease the surface tension. If water is the solvent, most organic substances show such a behaviour. The LiCl adsorption isotherm is characteristic for lyophilic substances. Most ions in water show such behaviour.

In order to describe the influence of a substance on the surface tension, one could specify the gradient of the adsorption isotherm for $c \to 0$. A list of these values for some substances dissolved in water at room temperature is shown in Table 3.2.

| Solute | $d(\Delta\gamma)/dc(10^{-3}\,\mathrm{Nm^{-1}M^{-1}})$ |
|---|---|
| HCl | $-0.28$ |
| LiCl | $1.81$ |
| NaCl | $1.82$ |
| CsCl | $1.54$ |
| $CH_3COOH$ | $-38$ |

**Table 3.2:** Gradient of the adsorption isotherm for $c \to 0$ of different solutes in water at 25°C.

■ **Example 3.4.** Adding 1 mM NaCl in water results in a slight increase of the surface tension of $\Delta\gamma = 1.82 \times 10^{-3}$ N/m $\times 0.001 = 1.82 \times 10^{-6}$ N/m. Upon addition of 1 mM $CH_3COOH$ the surface tension decreases to $3.8 \times 10^{-5}$ N/m.

### 3.4.4 The Marangoni effect

Often a typical shape of wine on the rim of a wine glass is observed, called tears of wine [51, 52]. To understand this phenomenon we take wine as a mixture of water and ethanol. A wine drop is thicker at the bottom than at the top because of its weight (Fig. 3.8). Part of the liquid evaporates. Due to its lower vapor pressure ethanol evaporates faster than water. The rate of evaporation is roughly proportional to the surface area. Since the drop is thinner at the

top, the concentration of ethanol decreases faster at the top. This causes the surface tension at the top to increase.

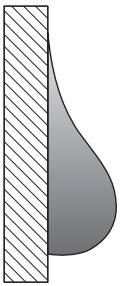


**Figure 3.8:** Marangoni effect: Tears of wine.

Tears of wine belong to a larger class of phenomena, which are all characterized by the carrying of bulk material through motions energized by surface tension gradients. It is called the Marangoni effect. Marangoni[3] observed this effect when studying the spreading of oil on large ponds of water [53].

## 3.5 Summary

- To apply the thermodynamic formalism to surfaces, Gibbs defined the ideal dividing plane which is infinitely thin. Excess quantities are defined with respect to a particular position of the dividing plane. The most important quantity is the interfacial excess which describes the amount of substance enriched or depleted at an interface.

- For a pure liquid the Gibbs dividing plane is conveniently positioned so that the surface excess is zero. Then the surface tension is equal to the surface free energy and the interfacial Gibbs free energy: $f^\sigma = g^\sigma = \gamma$.

- For solutions the Gibbs dividing plane is conveniently positioned so that the surface excess of the solvent is zero. Then the Gibbs adsorption isotherm (Eq. 3.52) relates the surface tension to the amount of solute adsorbed at the interface:

  $$\Gamma = -\frac{a}{RT} \cdot \frac{\partial \gamma}{\partial a}$$

  When the solute is enriched at the interface, the surface tension decreases upon addition of a solute. When a solute avoids the interface, the surface tension increases when adding the substance.

[3] Carlo Marangoni, 1840–1925. Italian physicist, professor at a Lyceum in Pavia.

## 3.6 Exercises

1. The surface tensions (in mN/m) of methanol and octane are

| | 10°C | 25°C | 50°C |
|---|---|---|---|
| Methanol | 23.23 | 22.07 | 20.14 |
| Octane | 22.57 | 21.14 | 18.77 |

   Calculate the surface entropy and the inner surface energy at 25°C.

2. Soap bubbles: To stabilize a bubble, surfactants are usually added to water. Assume we add a surfactant to a concentration of 2 mM. At this concentration we have a positive surface excess. As an average, each surfactant molecule occupies a surface area of 0.7 $nm^2$. Estimate the change in pressure inside a soap bubble with a radius of 1 cm compared to a hypothetical bubble formed from pure water.

3. Equation (3.52) can also be used to describe the adsorption of gases to surfaces. Then it can be written as

$$\Gamma = -\frac{1}{RT} \cdot \frac{d\gamma}{d \ln P}, \tag{3.55}$$

   with $P$ being the partial pressure of the adsorbing gas. Derive Eq. (3.55) from Eq. (3.52).

# 4 The electric double layer

## 4.1 Introduction

This chapter is about charged solid surfaces in liquids. The most important liquid is water. Because of its high dielectric constant water is a good solvent for ions. For this reason most surfaces in water are charged. Different processes can lead to charging. Alternatively, ions adsorb to a surface or dissociate from a surface. A protein might, for instance, expose an amino group on its surface. This can become protonated and thus positively charged ($\sim NH_2 + H^+ \rightarrow \sim NH_3^+$). Oxides are often negatively charged in water due to the dissociation of a proton from a surface hydroxyl group ($\sim OH \rightarrow \sim O^- + H^+$). Another way of charging a conducting surface is to apply an external electric potential between this surface and a counterelectrode. This is typically done in an electrochemical cell. How surfaces become charged and how this surface charge is quantitatively related to specific properties of the surface, is described in the next chapter.

Surface charges cause an electric field. This electric field attracts counter ions. The layer of surface charges and counter ions is called "electric double layer". In the simplest model of an electric double layer, the counter ions bind directly to the surface and neutralize the surface charges much like in a plate capacitor. In honor of the work of Ludwig Helmholtz[1] on electric capacitors it is called the Helmholtz layer. The electric field generated by the surface charges is accordingly limited to the thickness of a molecular layer. Based on the Helmholtz picture one could interpret some basic features of charged surfaces, but the model failed to explain one of the main properties which could be easily measured: the capacitance of an electric double layer.

In the years 1910–1917 Gouy[2] and Chapman[3] went a step further. They took into account a thermal motion of the ions. Thermal fluctuations tend to drive the counterions away form the surface. They lead to the formation of a diffuse layer, which is more extended than a molecular layer. For the simple case of a planar, negatively charged plane this is illustrated in Fig. 4.1. Gouy and Chapman applied their theory on the electric double layer to planar surfaces [54–56]. Later, Debye and Hückel calculated the potential and ion distribution around spherical surfaces [57].

[1] Hermann Ludwig Ferdinand Helmholtz, 1821–1894. German physicist and physiologist, professor in Königsberg, Bonn, Heidelberg, and Berlin.

[2] Louis George Gouy, 1854–1926. French physicist, professor in Lyon.

[3] David Leonard Chapman, 1869–1858. English chemist, professor in Manchester and Oxford.

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

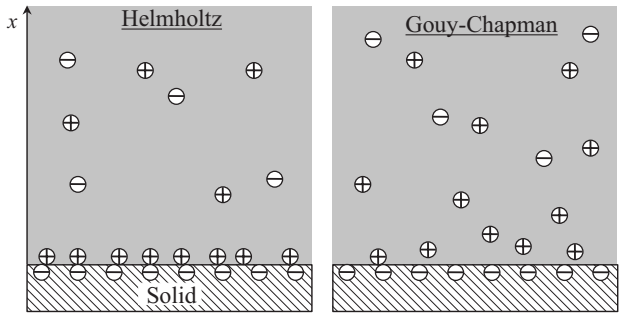



**Figure 4.1:** Helmholtz and Gouy–Chapman model of the electric double layer.

Both, the Gouy–Chapman and Debye–Hückel are continuum theories. They treat the solvent as a continuous medium with a certain dielectric constant, but they ignore the molecular nature of the liquid. Also the ions are not treated as individual point charges, but as a continuous charge distribution. For many applications this is sufficient and the predictions of continuum theory agree with experimental results. At the end of this chapter we discuss the limitations and problems of the continuum model.

## 4.2 Poisson–Boltzmann theory of the diffuse double layer

### 4.2.1 The Poisson–Boltzmann equation

The aim of this chapter is to calculate the electric potential $\psi$ near a charged planar interface. In general, this potential depends on the distance normal to the surface $x$. Therefore, we consider a planar solid surface with a homogeneously distributed electric surface charge density $\sigma$, which is in contact with a liquid. The surface charge generates a surface potential $\psi_0 = \psi(x = 0)$.

What is the potential distribution $\psi(x, y, z)$ in the solution? In general, charge density and electric potential are related by the Poisson[4] equation:

$$\nabla^2\psi = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} = -\frac{\rho_e}{\varepsilon\varepsilon_0} \tag{4.1}$$

Here, $\rho_e$ is the local electric charge density in C/m$^3$. With the Poisson equation, the potential distribution can be calculated once the exact charge distribution is known. The complication in our case is that the ions in solution are free to move. Before we can apply the Poisson equation we need to know more about their spacial distribution. This information is provided by Boltzmann[5] statistics. According to the Boltzmann equation the local ion density is given by

$$c_i = c_i^0 \cdot e^{-W_i/k_BT}, \tag{4.2}$$

[4] Denis Poisson, 1781–1840. French mathematician and physicist, professor in Paris.

[5] Ludwig Boltzmann, 1844–1906. Austrian physicist, professor in Vienna.

where $W_i$ is the work required to bring an ion in solution from infinite distance to a certain position closer to the surface. Eq. (4.2) tells us how the local ion concentration $c_i$ of the $i^{th}$ species depends on the electric potential at a certain position. For example, if the potential at a certain position in solution is positive, the chance of finding an anion at this position there is increased, while the cation concentration is reduced.

Now we assume that only electric work has to be done. We neglect for instance that the ion must displace other molecules. In addition, we assume that only a 1:1 salt is dissolved in the liquid. The electric work required to bring a charged cation to a place with potential $\psi$ is $W^+ = e\psi$. For an anion it is $W^- = -e\psi$. The local anion and cation concentrations $c^-$ and $c^+$ are related with the local potential $\psi$ by the Boltzmann factor: $c^- = c_0 \cdot e^{e\psi/k_BT}$ and $c^+ = c_0 \cdot e^{-e\psi/k_BT}$. Here, $c_0$ is the bulk concentration of the salt. The local charge density is

$$\rho_e = e(c^+ - c^-) = c_0 e \cdot \left(e^{-\frac{e\psi(x,y,z)}{k_BT}} - e^{\frac{e\psi(x,y,z)}{k_BT}}\right) \tag{4.3}$$

To remind you that the potential depends on the position we explicitly wrote $\psi(x, y, z)$. Substituting the charge density into the Poisson Eq. (4.1) leads to

$$\nabla^2\psi = \frac{c_0 e}{\varepsilon\varepsilon_0} \cdot \left(e^{\frac{e\psi(x,y,z)}{k_BT}} - e^{-\frac{e\psi(x,y,z)}{k_BT}}\right) \tag{4.4}$$

Often this equation is referred to as the Poisson–Boltzmann equation. It is a partial differential equation of second order, which in most cases has to be solved numerically. Only for some simple geometries can it be solved analytically. One such geometry is a planar surface.

### 4.2.2 Planar surfaces

For the simple case of a planar, infinitely extended planar surface, the potential cannot change in the $y$ and $z$ direction because of the symmetry and so the differential coefficients with respect to $y$ and $z$ must be zero. We are left with the Poisson–Boltzmann equation which contains only the coordinate normal to the plane $x$:

$$\frac{d^2\psi}{dx^2} = \frac{c_0 e}{\varepsilon\varepsilon_0} \cdot \left(e^{\frac{e\psi(x)}{k_BT}} - e^{-\frac{e\psi(x)}{k_BT}}\right) \tag{4.5}$$

Before we solve this equation for the general case, it is illustrative and for many applications, sufficient, to treat the special case of low potentials. How does the potential change with distance for low potentials? "Low" means, in a strict sense, $e\,|\psi| \ll k_BT$. At room temperature this is $\psi \leq 25$ mV. Fortunately, in most applications the result is valid even for higher potentials, up to 50–80 mV.

For low potentials we can expand the exponential functions into a series and neglect all but the first (i.e. the linear) term:

$$\frac{d^2\psi}{dx^2} = \frac{c_0 e}{\varepsilon\varepsilon_0} \cdot \left(1 + \frac{e\psi}{k_BT} - 1 + \frac{e\psi}{k_BT} \pm \ldots\right) \approx \frac{2c_0 e^2}{\varepsilon\varepsilon_0 k_BT} \cdot \psi \tag{4.6}$$

This is sometimes called the "linearized Poisson–Boltzmann equation". The general solution of the linearized Poisson–Boltzmann equation is

$$\psi(x) = C_1 \cdot e^{-\kappa x} + C_2 \cdot e^{\kappa x} \tag{4.7}$$

with

$$\kappa = \sqrt{\frac{2c_0e^2}{\varepsilon\varepsilon_0 k_B T}} \tag{4.8}$$

$C_1$ and $C_2$ are constants, which are defined by the boundary conditions. The boundary conditions require that, at the surface, the potential is equal to the surface potential, $\psi(x = 0) = \psi_0$, and that, for large distances from the surface, the potential should disappear $\psi(x \to \infty) = 0$. The second boundary condition guarantees that, for very large distances, the potential becomes zero and does not grow infinitely. It directly leads to $C_2 = 0$. From the first boundary condition we get $C_1 = \psi_0$. Hence, the potential is given by

$$\psi = \psi_0 \cdot e^{-\kappa x} \tag{4.9}$$

The potential decreases exponentially. The decay length is given by $\lambda_D = \kappa^{-1}$. It is called the **Debye**[6] **length**.

The Debye length decreases with increasing salt concentration. This is plausible, because the more ions are in solution, the more effective is the screening of surface charge. If we quantify all factors for water at 25°C, then for a monovalent salt, the Debye length is

$$\lambda_D = \frac{3.04\,\text{Å}}{\sqrt{c_0 \frac{\text{L}}{\text{mol}}}} \tag{4.10}$$

with the concentration $c_0$ in mol/L. For example, the Debye length of a 0.1 M aqueous NaCl solution at 25°C is 0.96 nm.

In water $\lambda_D$ cannot be longer than 680 nm. Due to the dissociation of water (according to $2\,H_2O \to H_3O^+ + OH^-$) the ion concentration cannot decrease below $2 \times 10^{-7}$ mol/L. Practically, the Debye length even in distilled water is only a few 100 nm due to ionic impurities or a pH different from pH 7.

Until now we have assumed that we are only dealing with monovalent so-called 1:1 salts. If ions of higher valency are also present, the inverse Debye length is given by

$$\kappa = \sqrt{\frac{e^2}{\varepsilon\varepsilon_0 k_B T} \sum_i c_i^0 Z_i^2} \tag{4.11}$$

Here, $Z_i$ is the valency of the $i^{th}$ ion sort. Please keep in mind: The concentrations have to be given in particles per $m^3$.

■ **Example 4.1.** Human blood plasma, that is blood without red and white blood cells and without thrombocytes, contains 143 mM $Na^+$, 5 mM $K^+$, 2.5 mM $Ca^{2+}$, 1 mM $Mg^{2+}$, 103 mM $Cl^-$, 27 mM $HCO_3^-$, 1 mM $HPO_4^{2-}$ and 0.5 mM $SO_4^{2-}$. What is the Debye length? We insert

[6] Peter Debye, 1884–1966. American physicist of Dutch origin, professor in Zürich, Utrecht, Göttingen, Leipzig, Berlin, and Ithaca. Nobel prize for chemistry, 1936.

| | | | |
|---|---|---|---|
| $c^0_{Na} = 861 \times 10^{23}\mathrm{m}^{-3}$ | $Z_{Na} = 1$ | $c^0_{K} = 30 \times 10^{23}\mathrm{m}^{-3}$ | $Z_{K} = 1$ |
| $c^0_{Ca} = 15 \times 10^{23}\mathrm{m}^{-3}$ | $Z_{Ca} = 2$ | $c^0_{Mg} = 6 \times 10^{23}\mathrm{m}^{-3}$ | $Z_{Mg} = 2$ |
| $c^0_{Cl} = 620 \times 10^{23}\mathrm{m}^{-3}$ | $Z_{Cl} = -1$ | $c^0_{HCO_3} = 163 \times 10^{23}\mathrm{m}^{-3}$ | $Z_{HCO_3} = -1$ |
| $c^0_{HPO_4} = 6 \times 10^{23}\mathrm{m}^{-3}$ | $Z_{HPO_4} = -2$ | $c^0_{SO_4} = 3 \times 10^{23}\mathrm{m}^{-3}$ | $Z_{SO_4} = -2$ |

into Eq. (4.11). Considering that at 36°C the dielectric constant of water is $\varepsilon = 74.5$ we get a Debye length of 0.78 nm.

All ions come from the dissociation of salts according to $AB \rightarrow A^- + B^+$. For this reason the sum $\sum c_i^\circ Z_i$ should always be zero (electroneutrality). Inserting the values from above, we find a surplus of cations of 22 mM. These cations come from the dissociation of organic acids (6 mM) and proteins (16 mM).

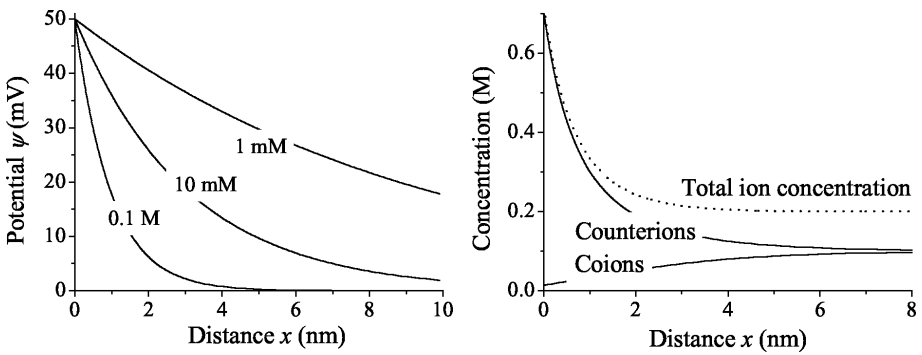



**Figure 4.2:** Left: Potential-versus-distance for a surface potential of $\psi_0$ = 50 mV and different concentrations of a monovalent salt in water. Right: Local co- and counterion concentrations are shown for a monovalent salt at a bulk concentration of 0.1 M and a surface potential of 50 mV. In addition, the total concentration of ions, that is the sum of the co- and counterion concentrations, is plotted.

Figure 4.2 illustrates several features of the diffuse electric double layer. The potential decreases exponentially with increasing distance. This decrease becomes steeper with increasing salt concentration. The concentration of co-ions is drastically increased close to the surface. As a result the total concentration of ions at the surface and thus the osmotic pressure is increased.

### 4.2.3 The full one-dimensional case

In many practical cases we can use the low-potential-assumption and it leads to realistic results. In addition, it is a simple equation and dependencies like the one on the salt concentration can easily be seen. In some cases, however, we have high potentials and we cannot linearize the Poisson–Boltzmann equation. Now we treat the general solution of the one-dimensional Poisson–Boltzmann equation and drop the assumption of low potentials. It is convenient to solve the equation with the dimensionless potential $y \equiv e\psi/k_BT$. Please do not mix this up with the spacial coordinate $y$! In this section we use the symbol "$y$" for the

dimensionless potential because many textbooks do so. The Poisson–Boltzmann equation for a 1:1 salt becomes

$$\frac{d^2y}{dx^2} = \frac{c_0 e^2}{\varepsilon\varepsilon_0 k_B T} \cdot \left(e^y - e^{-y}\right) = \frac{2c_0 e^2}{\varepsilon\varepsilon_0 k_B T} \cdot \frac{1}{2}\left(e^y - e^{-y}\right) = \kappa^2 \cdot \sinh y \tag{4.12}$$

Using $\sinh y = 1/2 \cdot (e^y - e^{-y})$. To solve the differential equation we multiply both sides by $2 \cdot dy/dx$:

$$2 \cdot \frac{dy}{dx} \cdot \frac{d^2y}{dx^2} = 2 \cdot \frac{dy}{dx} \cdot \kappa^2 \cdot \sinh y \tag{4.13}$$

The left side is equal to $\frac{d}{dx}\left(\frac{dy}{dx}\right)^2$. We insert this and integrate:

$$\int \frac{d}{dx'}\left(\frac{dy}{dx'}\right)^2 dx' = 2\kappa^2 \cdot \int \frac{dy}{dx'} \cdot \sinh y \cdot dx' \Leftrightarrow$$

$$\left(\frac{dy}{dx}\right)^2 = 2\kappa^2 \cdot \int \sinh y' \cdot dy' = 2\kappa^2 \cdot \cosh y + C_1 \tag{4.14}$$

$C_1$ is an integration constant. It is determined by the boundary conditions. At large distances the dimensionless potential $y$ and its derivative $dy/dx$ are zero. Since $\cosh y = 1$ for $y = 0$ this constant is $C_1 = -2\kappa^2$. It follows that

$$\left(\frac{dy}{dx}\right)^2 = 2\kappa^2 \cdot (\cosh y - 1) \Rightarrow \frac{dy}{dx} = -\kappa \cdot \sqrt{2\cosh y - 2} \tag{4.15}$$

In front of the square root there is a minus sign because $y$ has to decrease for a positive potential with increasing distance, i.e. $y > 0 \Rightarrow dy/dx < 0$. Now we remember the mathematical identity $\sinh \frac{y}{2} = \sqrt{\frac{1}{2}(\cosh y - 1)}$. Thus

$$\frac{dy}{dx} = -2\kappa \cdot \sinh \frac{y}{2} \tag{4.16}$$

Separation of variables and integration leads to

$$\frac{dy}{\sinh \frac{y}{2}} = -2\kappa \cdot dx \Rightarrow \int \frac{dy'}{\sinh \frac{y'}{2}} = -2\kappa \cdot \int dx' \Rightarrow$$

$$2 \cdot \ln\left(\tanh \frac{y}{4}\right) = -2\kappa x + 2C_2 \tag{4.17}$$

$C_2$ is another integration constant. Written explicitly we get

$$\ln\left(\frac{e^{y/4} - e^{-y/4}}{e^{y/4} + e^{-y/4}}\right) = -\kappa x + C_2 \tag{4.18}$$

Multiplying the denominator and numerator (in brackets) by $e^{y/4}$ leads to

$$\ln\left(\frac{e^{y/2}-1}{e^{y/2}+1}\right) = -\kappa x + C_2 \tag{4.19}$$

Using the dimensionless surface potential $y_0 = y(x=0) = e\psi_0/k_BT$ we can determine the integration constant

$$\ln\left(\frac{e^{y_0/2}-1}{e^{y_0/2}+1}\right) = C_2 \tag{4.20}$$

Substituting the results in Eq. (4.19)

$$\ln\left(\frac{e^{y/2}-1}{e^{y/2}+1}\right) - \ln\left(\frac{e^{y_0/2}-1}{e^{y_0/2}+1}\right) = \ln\left(\frac{e^{y/2}-1+e^{y_0/2}+1}{e^{y/2}+1+e^{y_0/2}-1}\right) = -\kappa x$$

$$\Rightarrow e^{-\kappa x} = \left(\frac{e^{y/2}-1+e^{y_0/2}+1}{e^{y/2}+1+e^{y_0/2}-1}\right) \tag{4.21}$$

Solving the equation for $e^{y/2}$ leads to the alternative expression

$$e^{y/2} = \frac{e^{y_0/2}+1+(e^{y_0/2}-1)\cdot e^{-\kappa x}}{e^{y_0/2}+1-(e^{y_0/2}-1)\cdot e^{-\kappa x}} \tag{4.22}$$

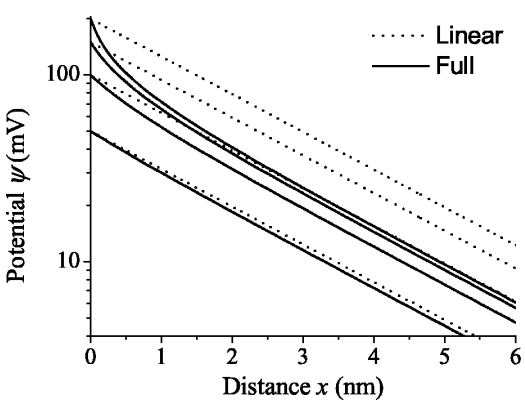



**Figure 4.3:** Potential-versus-distance for different surface potentials (50, 100, 150, and 200 mV) with 20 mM monovalent salt. The full solution Eq. (4.21) and the solution of the linearized Poisson–Boltzmann Eq. (4.9) are shown.

Let us compare results obtained with the linearized Poisson–Boltzmann Eq. (4.9) with the full solution Eq. (4.21). Figure 4.3 shows the potential calculated for a monovalent salt at a concentration of 20 mM in water. The Debye length is 2.15 nm. For a low surface potential of 50 mV both results agree well. When increasing the surface potential to 100, 150, or even 200 mV the full solution leads to lower potentials. At distances below $\approx \lambda/2$ the decay is therefore steeper than just the exponential decay. This steep decay at small distances becomes progressively more effective at higher and higher surface potentials, which leads to saturation behavior. For example, the potential at a distance of one Debye length can never exceed 40 mV irrespective of the surface potential. Such saturation behavior has indeed been observed experimentally [58].

### 4.2.4 The Grahame equation

In many cases we have an idea about the number of charged groups on a surface. Then we might want to know the potential. The question is: how are surface charge $\sigma$ and surface potential $\psi_0$ related? This question is also important because if we know $\sigma(\psi_0)$ we can calculate $d\sigma/d\psi_0$. This is basically the capacitance of the double layer and can be measured. The measured capacitance can be compared to the theoretical result to verify the whole theory.

Grahame derived an equation between $\sigma$ and $\psi_0$ based on the Gouy–Chapman theory. We can deduce the equation easily from the so-called electroneutrality condition. This condition demands that the total charge, i.e. the surface charge plus the charge of the ions in the whole double layer, must be zero. The total charge in the double layer is $\int_0^\infty \rho_e\, dx$ and we get [59]

$$\sigma = -\int_0^\infty \rho_e dx \tag{4.23}$$

Using the one-dimensional Poisson equation and the fact that at large distances the potential, and thus its gradient, are zero ($d\psi/dx|_{z=\infty} = 0$) we get

$$\sigma = \varepsilon\varepsilon_0 \int_0^\infty \frac{d^2\psi}{dx^2} \cdot dx = -\varepsilon\varepsilon_0 \left.\frac{d\psi}{dx}\right|_{x=0} \tag{4.24}$$

With $dy/dx = -2\kappa \cdot \sinh(y/2)$ and

$$\frac{dy}{dx} = \frac{d\left(e\psi/k_BT\right)}{dx} = \frac{e}{k_BT} \cdot \frac{d\psi}{dx} \tag{4.25}$$

we get the Grahame equation:

$$\sigma = \sqrt{8c_0\varepsilon\varepsilon_0 k_BT} \cdot \sinh\left(\frac{e\psi_0}{2k_BT}\right) \tag{4.26}$$

For low potentials we can expand sinh into a series ($\sinh x = x + x^3/3! + ...$) and ignore all but the first term. That leads to the simple relationship

$$\sigma = \frac{\varepsilon\varepsilon_0\psi_0}{\lambda_D} \tag{4.27}$$

■ **Example 4.2.** On the surface of a certain material there is one ionized group per (4 nm)$^2$ in aqueous solution containing 10 mM NaCl. What is the surface potential?

With a Debye length of 3.04 nm at 25°C the surface charge density in SI units is $\sigma = 1.60 \times 10^{-19}\mathrm{As}/16 \times 10^{-18}\mathrm{m}^2 = 0.01\,\mathrm{Asm}^{-2}$. With this we get

$$\psi_0 = \frac{\sigma\lambda_D}{\varepsilon\varepsilon_0} = \frac{0.01\,\mathrm{Asm}^{-2} \cdot 3.04 \times 10^{-9}\mathrm{m}}{78.4 \cdot 8.85 \times 10^{-12}\mathrm{AsV}^{-1}\mathrm{m}^{-1}} = 0.0438\,\mathrm{V} \tag{4.28}$$

Using the Grahame equation (4.26) we get a surface potential of 39.7 mV.

Figure 4.4 shows the calculated relationship between the surface potential and surface charge for different concentrations of a monovalent salt. We see that for small potentials the surface charge density is proportional to the surface potential. Depending on the salt concentration, the linear approximation (dashed) is valid till $\psi_0 \approx 40...80$ mV. At high salt concentration more surface charge is required to reach the same surface potential than for a low salt concentration.

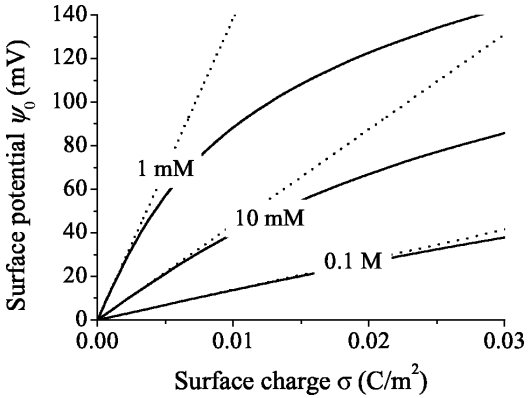



**Figure 4.4:** Surface potential-versus-surface charge calculated with the full Grahame equation (4.26, continuous line) and with the linearized version Eq. (4.27, dotted).

### 4.2.5 Capacity of the diffuse electric double layer

The differential capacitance between two regions of separated charges is, in general, defined as $dQ/dU$. Here, $Q$ is the charge on each "electrode" and $U$ is the voltage. The capacity of an electric double layer per unit area is thus

$$C_{GC}^{A} = \frac{d\sigma}{d\psi_0} = \sqrt{\frac{2e^2 c_0 \varepsilon\varepsilon_0}{k_B T}} \cdot \cosh\left(\frac{e\psi_0}{2k_B T}\right) = \frac{\varepsilon\varepsilon_0}{\lambda_D} \cdot \cosh\left(\frac{e\psi_0}{2k_B T}\right) \tag{4.29}$$

The index "$GC$" is a reminder that we calculated the capacitance in the Gouy–Chapman model. We expand cosh into a series ($\cosh x = 1 + x^2/2! + x^4/4! + ...$) and consider only the first term. Then we get

$$C_{GC}^{A} = \frac{\varepsilon\varepsilon_0}{\lambda_D} \tag{4.30}$$

It is instructive to compare this to the capacitance of a plate capacitor $\varepsilon\varepsilon_0 A/d$. Here, $A$ is the cross-sectional area and $d$ is the separation between the two plates. We see that the electric double layer behaves like a plate capacitor, in which the distance between the plates is given by the Debye length! The capacity of a double layer — that is the ability to store charge — rises with increasing salt concentration because the Debye length decreases.

To avoid confusion we should point out that $C_{GC}^{A}$ as defined above is the differential capacitance. The integral capacitance per unit area is $\sigma/\psi_0$. Experimentally the differential capacitance is easier to measure.

## 4.3 Beyond Poisson–Boltzmann theory

### 4.3.1 Limitations of the Poisson–Boltzmann theory

In the treatment of the diffuse electric double layer, several assumptions were made which lead to imperfections [60]:

- The finite size of the ions was neglected. In particular, close to the surface, this is a daring assumption because the ion concentration can get very high. For example, if we have a surface potential of 100 mV the counterion concentration is increased by a factor $\approx 50$. At a bulk concentration of 0.1 M the Poisson–Boltzmann theory predicts a local concentration at the surface of roughly 5 M. Then the ions have an average distance less than 1 nm. Considering the diameter of an ion with its hydration shell of typically 3–6 Å the detailed molecular structure should become significant.

- Ions in solution were considered as a continuous charge distribution. We ignored their discrete nature, namely that they can only carry a multiple of the unit charge. Also the surface charge is assumed to be homogeneous and smeared out. In reality it is formed by individual adsorbed ions or charged groups.

- All non-Coulombic interactions were disregarded. In water, for example, each ion has a hydration shell. If the ions approach each other, these hydration shells overlap and change the interaction drastically.

- The solvent is supposed to be continuous and we took the permittivity of the medium to be constant. This is certainly a rough approximation because polar molecules are hindered from rotating freely in the strong electric field at the surface. In addition, the high concentration of counterions in the proximity of the surface can change the permittivity drastically.

- Surfaces are assumed to be flat on the molecular scale. In many cases this is not a reasonable assumption. If we, for instance, consider a biological membrane in a physiological buffer the ion concentration is roughly 150 mM leading to a Debye length of 0.8 nm. Charges in phospholipids are distributed over a depth of up to 8 Å. In addition, the molecules thermally jump up and down so that the charges are distributed over a depth of almost 1 nm.

- Image forces between the ions and the surface were ignored.

Despite these strong assumptions, the Poisson–Boltzmann theory describes electric double layers surprisingly well. The reason is that errors lead to opposite effects and compensate each other. Including non-coulombic interactions leads to an increase of the ion concentration at the surface and a reduced surface potential. On the other hand, taking the finite size of the ions into account leads to a lower ion concentration at the surface and thus an increased surface potential. A reduction of the dielectric permittivity due to the electric field increases its range but at the same time reduces the surface potential because less ions dissociate or adsorb.

In summary: For aqueous solutions the Gouy–Chapman theory provides relatively good predictions for monovalent salts at concentrations below 0.2 M and for potentials below 50–80 mV. The fact that the surface charge in reality is not continuously but discretely distributed, leads, according to experience, to deviations only with bivalent and trivalent charges. Often, however, the surface charges do not lie precisely in one plane. This is true, for example, for biological membranes. In this case larger deviations might result.

How can we improve the Poisson–Boltzmann theory? Basically it is not that difficult to account for one or two of the mentioned defects but such improvements have limited practical relevance. Since the defects introduced by the assumptions lead to compensating errors, removing a defect might even lead to less realistic results. The most rigorous approach to improve the Poisson–Boltzmann theory is by starting from first principles, applying statistical thermodynamic equations for bulk electrolytes and non-uniform fluids. Excellent reviews about the statistical mechanics of double layers have appeared [61,62]. Also computer simulations increased our understanding of electric double layers significantly [63]. The significance of statistical mechanics and computer simulations remains, however, somewhat academic because it does not lead to simple analytical formulas that can easily be applied. In addition, for a real complete description, the real molecular nature of the solvent should be considered.

Fortunately, there is a relatively simple semi-empirical extension of the Gouy–Chapman theory, which accounts for most experimental observations. This extension was proposed by Stern.[7]

### 4.3.2 The Stern layer

Stern combined the ideas of Helmholtz and that of a diffuse layer [64]. In Stern theory we take a pragmatic, though somewhat artificial, approach and divide the double layer into two parts: an inner part, the Stern layer, and an outer part, the Gouy or diffuse layer. Essentially the Stern layer is a layer of ions which is directly adsorbed to the surface and which is immobile. In contrast, the Gouy–Chapman layer consists of mobile ions, which obey Poisson–Boltzmann statistics. The potential at the point where the bound Stern layer ends and the mobile diffuse layer begins is the zeta potential ($\zeta$ potential). The zeta potential will be discussed in detail in Section 5.4.

Stern layers can be introduced at different levels of sophistication. In the simplest case we only consider the finite size effect of the counterions (Fig. 4.5). Due to their size, which in water might include their hydration shell, they cannot get infinitely close to the surface but always remain at a certain distance. This distance $\delta$ between the surface and the centers of these counterions marks the so-called outer Helmholtz plane. It separates the Stern from the Gouy–Chapman layer. For a positively charged surface this is indicated in Fig. 4.5.

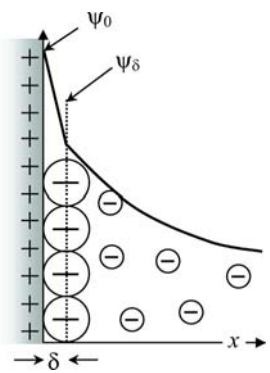



**Figure 4.5:** Simple version of the Stern layer.

[7] Otto Stern, 1888–1969. German physicist, professor in Hamburg. Nobel Prize in physics in 1943.

At the next level we also take specific adsorption of ions into account (Fig. 4.6). Specifically adsorbed ions bind tightly at a short distance. This distance characterizes the inner Helmholtz plane. In reality all models can only describe certain aspects of the electric double layer. A good model for the structure of many metallic surfaces in an aqueous medium is shown in Fig. 4.6. The metal itself is negatively charged. This can be due to an applied potential or due to the dissolution of metal cations. Often anions bind relatively strongly, and with a certain specificity, to metal surfaces. Water molecules show a distinct preferential orientation and thus a strongly reduced permittivity. They determine the inner Helmholtz plane.

Next comes a layer of nonspecifically adsorbed counterions with their hydration shell. Still, the permittivity is significantly reduced because the water molecules are not free to rotate. This layer specifies the outer Helmholtz plane. Finally there is the diffuse layer. A detailed discussion of the structure of the electric double layer at a metal surface is included in Ref. [65].

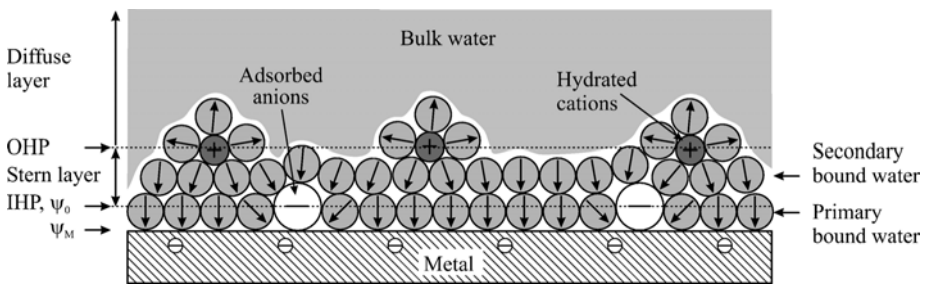



**Figure 4.6:** Stern layer at a metal surface. Due to the high electrical conductivity, the potential in the metal $\psi_M$ is constant up to the surface. The inner (IHP ) and outer (OHP) Helmholtz planes are indicated. In the first layer of primary bound water the permittivity is typically $\varepsilon = 6$. In the secondary layer of water it is of the order of $\varepsilon \approx 30$.

An important quantity with respect to experimental verification is the differential capacitance of the total electric double layer. In the Stern picture it is composed of two capacitors in series: the capacity of the Stern layer, $C^A_{St}$, and the capacitance of the diffuse Gouy–Chapman layer. The total capacitance per unit area is given by

$$\frac{1}{C^A} = \frac{1}{C^A_{St}} + \frac{1}{C^A_{GC}} \tag{4.31}$$

Let us estimate $C_{St}$ using the simple equation for a plate capacitor. The two plates are formed by the surface and by the adsorbed ions. Denoting the radius of the hydrated ions by $R_{ion}$ the distance is in the order of $R_{ion}/2 \approx 2$ Å. The capacitance per unit area of the Stern layer is $C^A_{St} = 2\varepsilon_{St}\varepsilon_0/r_{ion}$. The permittivity at the surface is reduced and typically of the order of $\varepsilon_{St} \approx 6 \ldots 32$ for water. Using a value of $\varepsilon_{St} = 10$ we estimate a capacitance for the Stern layer of $C^A_{St}$ =0.44 Fm$^{-2}$ = 44 $\mu$Fcm$^{-2}$. Experimental values are typically 10–100 $\mu$Fcm$^{-2}$.

## 4.4 The Gibbs free energy of the electric double layer

What is the Gibbs free energy of an electric double layer? The energy of an electric double layer plays a central role in colloid science, for instance to describe the properties of charged polymers (polyelectrolytes) or the interaction between colloidal particles. Here, we only give results for diffuse layers because it is simpler and in most applications only the diffuse layer is relevant. The formalism is, however, applicable to other double layers as well.

In order to calculate the Gibbs free energy of a Gouy–Chapman layer we split its formation into three steps [66, 67]. In reality it is not possible to do these steps separately but we can do the Gedanken experiment without violating any physical principle.

First, the uncharged colloidal particle is brought into an infinitely large solution. This solution contains specifically adsorbing ions and indifferent ions. Surface ions will bind or dissociate from surface groups, driven by chemical forces. It is important to realize that these chemical forces drive the formation of the double layer. To calculate the chemical energy we have to realize that the dissociation (or binding) of ions does not proceed forever because the more ions that dissociate the higher the electric potential becomes. This potential prevents further ions from dissociating from the surface. The process stops when the chemical energy is equal to the electrostatic energy. The electrostatic energy of one ion of charge $Q$ at this point is simply $Q\psi_0$. The chemical energy of this ion is $-Q\psi_0$. Hence, the Gibbs free energy per unit area for the formation of an electric double layer is $-\sigma\psi_0$.

Second, we bring the counterions to the surface. They are supposed to go directly to the surface at $x = 0$. The number of counterions is equal to the number of ions in the diffuse double layer. The first counterions are still attracted by the full surface potential. Their presence, however, reduces the surface potential and the following counterions only notice a reduced surface potential. To bring counterions to the surface the work $dG = \psi_0' d\sigma$ has to be performed, where $\psi_0'$ is the surface potential at a certain time of charging — or better discharging — process. The total energy we gain (and the double layer loses) is

$$\int_0^\sigma \psi_0' d\sigma'. \tag{4.32}$$

In a third step the counterions are released from the surface. Stimulated by thermal fluctuations, they partially diffuse away from the surface and form the diffuse double layer. The entropy and, at the same time, the energy increases. One can show that both terms compensate, so that in the third step no contribution to the Gibbs free energy results.

Summing up all contributions we obtain the total Gibbs free energy of the diffuse double layer per unit area:

$$g = -\sigma\psi_0 + \int_0^\sigma \psi_0' d\sigma' \tag{4.33}$$

In mathematics we learnt that

$$d\left(\psi_0'\sigma'\right) = \sigma' d\psi_0' + \psi_0' d\sigma' \Rightarrow \int d\left(\psi_0'\sigma'\right) = \int \sigma' d\psi_0' + \int \psi_0' d\sigma' \tag{4.34}$$

Using this equation we can write:

$$g = -\sigma\psi_0 + \int_0^{\sigma\psi_0} d(\sigma'\psi_0') - \int_0^{\psi_0} \sigma' d\psi_0' = -\int_0^{\psi_0} \sigma' d\psi_0' \tag{4.35}$$

The integral can be solved with the help of Grahames equation (4.26):

$$\begin{aligned} g &= -\int_0^{\psi_0} \sigma d\psi_0' = -\int_0^{\psi_0} \sqrt{8c_0\varepsilon\varepsilon_0 k_B T} \cdot \sinh\left(\frac{e\psi_0'}{2k_B T}\right) \cdot d\psi_0' \\ &= -\sqrt{8c_0\varepsilon\varepsilon_0 k_B T} \cdot \frac{2k_B T}{e} \cdot \left[\cosh\left(\frac{e\psi_0'}{2k_B T}\right)\right]_0^{\psi_0} \\ &= -8c_0 k_B T \lambda_D \cdot \left[\cosh\left(\frac{e\psi_0}{2k_B T}\right) - 1\right] \end{aligned} \tag{4.36}$$

For low potentials we can use the, even simpler, relation (4.27) and get

$$g = -\kappa\varepsilon\varepsilon_0 \cdot \int_0^{\psi_0} \psi_0' d\psi_0' = -\frac{\kappa\varepsilon\varepsilon_0}{2}\psi_0^2 = -\frac{1}{2}\sigma\psi_0 \tag{4.37}$$

The Gibbs free energy of an electric double layer is negative because it forms spontaneously. Roughly, it increases in proportion to the square of the surface potential.

■ **Example 4.3.** Estimate the energy per unit area of an electric double layer for a surface potential of 40 mV in an aqueous solution containing 0.01 M monovalent ions. Estimate the total energy of a colloidal particle of 3 nm diameter.

With

$$g = -\frac{\kappa\varepsilon\varepsilon_0}{2}\psi_0^2 = -\frac{\varepsilon\varepsilon_0}{2\lambda_D}\psi_0^2$$

and a Debye length of 3.04 nm we get

$$g = -\frac{78.4 \cdot 8.85 \times 10^{-12}\mathrm{AsV^{-1}m^{-1}}}{2 \cdot 3.04 \times 10^{-9}\mathrm{m}} \cdot (0.04\,\mathrm{V})^2 = -0.183 \times 10^{-3}\frac{\mathrm{J}}{\mathrm{m}^2}$$

Compared to typical surface tensions of liquids, this is small.

## 4.5 Summary

- Most solid surfaces in water are charged. Reason: Due to the high dielectric permittivity of water, ions are easily dissolved. The resulting electric double layer consist of an inner Stern or Helmholtz layer, which is in close contact with the solid surface, and a diffuse layer, also called the Gouy–Chapman layer.
- The electric potential in the diffuse layer of a planar surface decays exponentially

  $$\psi = \psi_0 \cdot e^{-x/\lambda_D},$$

  provided the potential does not exceed 50–80 mV.
- For a monovalent salt and at 25°C the Debye length is given by $\lambda_D = 3.04/\sqrt{c_0}$ Å, with the concentration $c_0$ in mol/L.

## 4.6 Exercises

1. Compare the Debye length of 0.1 mM NaCl solution of water and ethanol ($\varepsilon = 25.3$).

2. Plot the potential versus distance for surface potentials of 60 mV, 100 mV, and 140 mV using the solution of the linearized and the full Poisson–Boltzmann equation for an aqueous solution with 2 mM KCl.

3. For an electrophysiological experiment you form an electrode from a 5 cm long platinum wire (0.4 mm diameter) by bending it in the shape of a spiral. Calculate the total capacitance of the diffuse electric double layer for aqueous solutions of a monovalent salt at concentrations of 0.1 and 0.001 M. Assume a low surface potential.

4. Silicon oxide has a typical surface potential in an aqueous medium of $-70$ mV in 50 mM NaCl at pH 9. Which concentration of cations do you roughly expect close to the surface? What is the average distance between two adjacent cations? What is the local pH at the surface?

5. The differential capacitance of a mercury electrode in an aqueous medium containing NaF has been measured at the point of zero charge. It is 6.0 $\mu$F/cm$^2$ at 1 mM, 13.1 $\mu$F/cm$^2$ at 10 mM, 20.7 $\mu$F/cm$^2$ at 100 mM, and 25.7 $\mu$F/cm$^2$ at 1 M concentration. Compare this with the result of the Gouy–Chapman theory and draw conclusions.

# 5 Effects at charged interfaces

In the last chapter we learned how to describe electric double layers mathematically. In this chapter we focus on the question, how surfaces become charged. This is obviously an important question because surface charges are the reason for the formation of electric double layers. In addition, we discuss experimental aspects of charged interfaces. This includes techniques to analyze the properties of electric double layers but also applications and effects of charged surfaces.

There are various techniques to measure different properties of electric double layers. A wide range of information was obtained from electrocapillary experiments. In an electrocapillary experiment the surface tension versus potential of a metallic surface is measured. From this the capacitance and the surface charge can be calculated. For technical reasons this is routinely only possible for mercury.

Surface charge density and surface potential are of primary interest. For insulating surfaces, charge can be determined by potentiometric or conductometric titration, though this is a tedious procedure. For metals, the relationship between surface charge and potential can be determined by measuring the capacitance. Finally we discuss electrokinetic effects. Electrokinetic experiments yield the potential at the outer Helmholtz plane.

In addition, there are techniques developed in other fields of colloid science, which are not directly related to classical electrochemistry. In surface force experiments, for instance, the distance dependence of the electric double layer is measured precisely. This will be discussed later.

## 5.1 Electrocapillarity

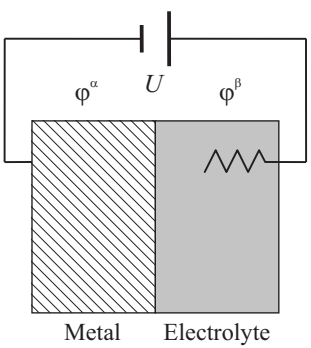



**Figure 5.1:** Schematic of a metal–electrolyte interface with an applied potential.

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

We start this chapter with electrocapillarity because it provides detailed information of the electric double layer. In a classical electrocapillary experiment the change of interfacial tension at a metal–electrolyte interface is determined upon variation of an applied potential (Fig. 5.1). It was known for a long time that the shape of a mercury drop which is in contact with an electrolyte depends on the electric potential. Lippmann[1] examined this electrocapillary effect in 1875 for the first time [68]. He succeeded in calculating the interfacial tension as a function of applied potential and he measured it with mercury.

### 5.1.1 Theory

The change of the interfacial tension can be calculated with the help of the Gibbs–Duhem equation even when a potential is applied. In order to use the equation, we first need to find out which molecular species are present. Evidently, only those which are free to move are of interest. In the electrolyte we have the dissolved ions. In the metal the electrons can move and have to be considered.

Since in addition to the chemical potentials $\mu_i$, also the electrical potential $\varphi$, affects the charged species, electrochemical potentials $\mu_i^*$ must be used. We use the symbol $\varphi$ instead of $\psi$ because this is the Galvani potential (see Section 5.5). The Gibbs–Duhem equation for changes of state functions at constant temperature is

$$d\gamma = -\sum_{i=1}^{n} \Gamma_i d\mu_i^* - \Gamma_e d\mu_e^* \tag{5.1}$$

where

$$d\mu_i^* = d\mu_i + Z_i F_A d\varphi \quad \text{and} \quad d\mu_e^* = d\mu_e - F_A d\varphi$$

The first term refers to the electrolyte. Accordingly, the sum runs over all ion types present in the electrolyte. The second term contains the contribution of the electrons in the metal. $\Gamma_i$ and $\Gamma_e$ are the interfacial excess concentrations of the ions in solution and of the electrons in the metal, respectively. $\mu_i$ is the chemical potential of the particle type $i$, $F_A$ is Faradays constant, and $\mu_e^*$ is the electrochemical potential of the electrons. Substitution leads to

$$d\gamma = -\sum_{i=1}^{n} \Gamma_i d\mu_i - F_A \sum_{i=1}^{n} \Gamma_i Z_i d\varphi^\beta - \Gamma_e d\mu_e + F_A \Gamma_e d\varphi^\alpha \tag{5.2}$$

Here, $\varphi^\alpha$ and $\varphi^\beta$ are the electrical (Galvani) potentials in the two phases.

What are the correct values of the potentials? In the metal the potential is the same everywhere and therefore $\varphi^\alpha$ has one clearly defined value. In the electrolyte, the potential close to the surface depends on the distance. Directly at the surface it is different from the potential one Debye length away from it. Only at a large distance away from the surface is the potential constant. In contrast to the electric potential, the electro*chemical*potential is the same everywhere in the liquid phase assuming that the system is in equilibrium. For this reason we use the potential and chemical potential far away from the interface.

[1] Gabriel Lippmann, 1845–1921. French physicist, Nobel Price in physics 1908.

The concentrations of ions and electrons are not independent. Assuming that the whole system is electrically neutral (electroneutrality) one degree of freedom is lost. It is convenient to regard the electrons in the metal as the dependent component. With the help of the electroneutrality condition $\sum \Gamma_i z_i = \Gamma_e$, the last term in Eq. (5.2) can be written as $F_A \sum \Gamma_i Z_i d\varphi^\alpha$. Then it is possible to summarize the second and fourth term and we get:

$$\begin{aligned} d\gamma &= -\sum_{i=1}^{n} \Gamma_i d\mu_i - F_A \sum_{i=1}^{n} \Gamma_i Z_i d\varphi^\beta - \Gamma_e d\mu_e + F_A \sum_{i=1}^{n} \Gamma_i Z_i d\varphi^\alpha \\ &= -\sum_{i=1}^{n} \Gamma_i d\mu_i - \Gamma_e d\mu_e - \sigma \cdot d(\varphi^\beta - \varphi^\alpha) \end{aligned} \tag{5.3}$$

This is the fundamental equation for the description of electrocapillarity. Thereby $\sigma = F_A \sum \Gamma_i Z_i$ is identified with the surface charge density, which is produced by electrons in the metal and compensated by ions in solution. This identification is generally doubtful, because the surface excesses $\Gamma_i$ depend on the position of the interface. If, however, the electrode is totally polarisable (no electrons are exchanged between the metal and the electrolyte), then the positioning of the interface is trivial and $\sigma$ represents the surface charge density.

The potential difference $\varphi^\beta - \varphi^\alpha$ is not necessarily equal to the externally applied voltage $U$. Usually both differ by a constant. A possible voltage drop in the electrolyte, in particular with low salt concentrations, should be considered too. Since $\varphi^\beta - \varphi^\alpha$ and $U$ differ only by a constant and $d(\varphi^\beta - \varphi^\alpha) = dU$, Eq. (5.3) is still valid. From the last expression the Lippman equation

$$\frac{\partial \gamma}{\partial U} = -\sigma \tag{5.4}$$

and the differential capacitance of the electrical double layer per unit area

$$\frac{\partial^2 \gamma}{\partial U^2} = -\frac{\partial \sigma}{\partial U} = C^A \tag{5.5}$$

are obtained directly .

In the derivation we assumed that no electrochemical reactions take place, i.e., the electrode should be totally polarisable. A more general derivation of the Lippmann equation, which explicitly considers electrochemical reactions, is found in many electrochemistry text books or Refs. [59, 69].

Lippmanns classical derivation was simpler. He treated the mercury–electrolyte interface as a capacitor. The capacitance is assumed to increase with the surface area $A$ much like a plate capacitor. One "plate" is the metal, the other the layer of counterions in the electrolyte. The potential difference between the two plates is $U$. A change in the Gibbs free energy of the system is equal to the reversible work upon a change of the surface $A$ or of the charge $Q$:

$$dG = \gamma dA + U dQ \tag{5.6}$$

The second term indicates the work required to bring $dQ$ from one plate to the other at a potential $U$. Assuming we have a homogeneous system, an integration at constant intrinsic parameters $\gamma$ and $U$ results in

$$G = \gamma A + UQ \tag{5.7}$$

A comparison of the total differential ($dG = \gamma dA + Ad\gamma + UdQ + QdU$) with Eq. (5.6) shows that

$$0 = Ad\gamma + QdU \Rightarrow \quad \frac{d\gamma}{dU} = -\frac{Q}{A} = -\sigma \tag{5.8}$$

The simplest electrocapillary curves are obtained under the assumption that $C^A$ is constant (see Fig. 5.2). Then we have

$$\gamma = \gamma_0 - \frac{1}{2}C^A U^2 \tag{5.9}$$

because

$$\gamma - \gamma_0 = \int_0^U \frac{d\gamma}{dU'}dU' = -\int_0^U \frac{Q}{A}dU' = -\int_0^U C^A U' dU' = -\frac{C^A U^2}{2} \tag{5.10}$$

Here, $\gamma_0 = \gamma(U = 0)$ is the interfacial tension at a potential where the interface is uncharged and the potential accordingly disappears.

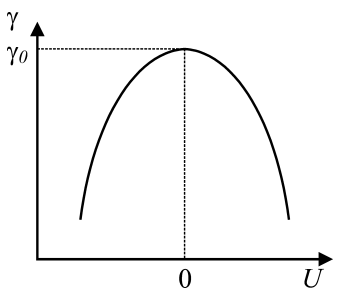



**Figure 5.2:** Schematic electrocapillary curve.

The interfacial tension decreases with increasing amount of surface potential. The reason is the increased interfacial excess of counterions in the electric double layer. In accordance with the Gibbs adsorption isotherms, the interfacial tension must decrease with increasing interfacial excess. At charged interfaces ions have an effect similarly to surfactants at liquid surfaces.

### 5.1.2 Measurement of electrocapillarity

Fundamental knowledge about the behavior of charged surfaces comes from experiments with mercury. How can an electrocapillarity curve of mercury be measured? A usual arrangement, the so-called dropping mercury electrode, is shown in Fig. 5.2 [70]. A capillary filled with mercury and a counter electrode are placed into an electrolyte solution. A voltage is applied between both. The surface tension of mercury is determined by the maximum bubble pressure method. Mercury is thereby pressed into the electrolyte solution under constant pressure $P$. The number of drops per unit time is measured as a function of the applied voltage.

There are different methods of measuring the electrocapillarity of solid metals. In one method the deflection of a cantilever spring, which is coated on one side with the metal, is measured [71, 72]. The surface tension of the metallic side changes when the applied voltage

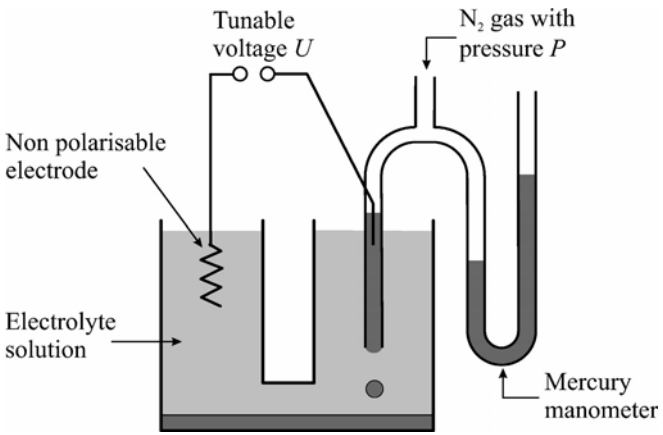



**Figure 5.3:** Dropping mercury electrode.

varies, which leads to a bending of the cantilever. If we, for instance, assume that the surface tension of the metallised side decreases, then the spring bends to the other side, and vice versa. In an alternative method we detect the expansion or the compression, of a strained wire under the effect of the surface tension [73, 74]. Finally, also a change of contact angle with variation in the potential can be measured. It contains information about the surface tension of the solid [75–77].

Unfortunately, all these methods are nowhere near as precise and easy to carry out as with mercury. The fundamental reason is that molecules in the solid are not mobile and the formation of a new surface is different from in a liquid. Thus also the theoretical treatment is substantially more difficult [78–80].

## 5.2 Examples of charged surfaces

In this section we discuss five different materials as examples with different charging mechanisms: mercury, silver iodide, oxides, mica, and semiconductors. Mercury is one example of an inert metal. Silver iodide is an example of a weakly soluble salt. Oxides are an important class of minerals. For most biological substances like proteins or lipids a similar charging process dominates. Mica is an example for a clay mineral. In addition, it is widely used as a substrate in surface force measurements and microscopy. We also included a general discussion of semiconductors because the potential in the semiconductor can be described similarly to the diffuse layer in electrolytes and there is an increasing effort to make a direct contact between a liquid or a living cell and a semiconductor.

The presented mechanisms are not the only ones occurring. Air bubbles or oil drops in water, for instance, are negatively charged, probably due to an adsorption of hydroxyl ions. The process is far from being understood. Many polymer surfaces acquire a negative surface charge in water. This could be due to anions, which are adsorbed due to the van der Waals force. Again, this process is not well understood.

### 5.2.1 Mercury

The mercury surface is probably the best characterized surface with respect to its electric properties. The explanation is that mercury is one of the few metals that is liquid at room temperature. Since it is a metal a voltage can easily be applied. Since it is a liquid the surface tension can be measured simply and precisely. Then the surface charge can be calculated with the help of the Lippmann equation. Additionally, a fresh surface free of contamination can be continuously produced.

If we measure electrocapillary curves of mercury in an aqueous medium which contains KF, NaF, or CsF, then we observe that the typical parabolas become narrower with increasing concentration. Explanation: With increasing salt concentration the Debye-length becomes shorter, the capacity of the double layer increases. The maximum of the electrocapillarity curve, and thus the point of zero charge (pzc), remains constant, i.e., neither the cations nor fluoride adsorb strongly to mercury.

Another behavior is observed in solutions with KOH, KCNS, and KI (Fig. 5.4). An increase in the concentration leads to a shift of the maximum to negative potentials. Explanation: The anions bind specifically to mercury and shift the point of zero charge. Iodide adsorbs more strongly than cyanide, and cyanide more strongly than hydroxide. A negative potential must be applied in order to drive the anions away from the surface.

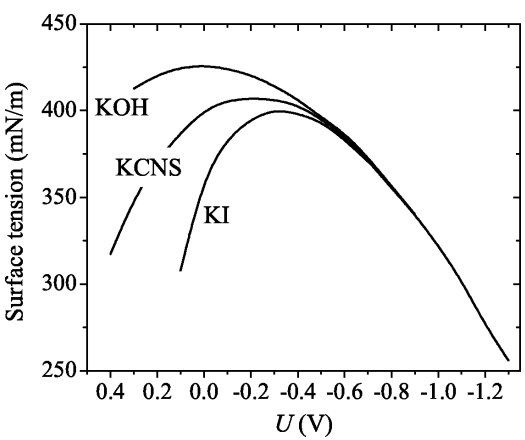



**Figure 5.4:** Electrocapillary curves of mercury measured in different aqueous electrolytes at 18°C. The zero of the applied electric potential was chosen to be at the maximum of the electrocapillary curve for electrolytes such as NaF, $Na_2SO_4$, and $KNO_3$, which do not strongly adsorb to mercury. Redrawn after Ref. [59].

Anions bind also to other metals, like gold, platinum, or silver [74, 81]. Why do anions adsorb specifically to metals, while cations do not? The explanation is a strong hydration of cations. A cation would have to give up its hydration shell for an adsorption. This is energetically disadvantageous. Anions are barely hydrated and can therefore bind more easily to metals [82]. Another possible explanation is the stronger van der Waals force between anions and metals. The binding of ions to metallic surfaces is not yet understood and even the idea that cations are not directly bound to the metal, was questioned based on molecular-dynamics simulations [83].

### 5.2.2 Silver iodide

If AgI or AgCl is submerged in water, a certain number of molecules dissolve (review: Ref. [84]). An equilibrium between the ions in solution and the crystal is established:

$$\mathrm{AgI} \rightleftharpoons \mathrm{Ag}^+ + \mathrm{I}^- \tag{5.11}$$

The concentrations of $\mathrm{Ag}^+$ and $\mathrm{I}^-$ in the solution are low because the solubility product $K = a_{Ag^+} \cdot a_{I^-} \approx 10^{-16}\,\mathrm{M}^2$ is small. Here, $a_{Ag^+}$ and $a_{I^-}$ are the activities of $\mathrm{Ag}^+$ and $\mathrm{I}^-$ in water, respectively.

The surface of AgI can be imagined as a regular lattice of $\mathrm{Ag}^+$ and $\mathrm{I}^-$ ions. With the same number of $\mathrm{Ag}^+$ and $\mathrm{I}^-$, the surface would be uncharged. It is, however, clear that this does not agree with the case where equal ion concentrations are present also in solution. Iodide has a somewhat higher affinity for the AgI surface than $\mathrm{Ag}^+$. In the case $[Ag^+] = [I^-] = 10^{-8}\mathrm{M}$, silver iodide is negatively charged. The surface can be neutralized by increasing the $\mathrm{Ag}^+$ concentration in solution, e.g. by addition of $\mathrm{AgNO_3}$. Already a small concentration of $10^{-5.5}$ M $\mathrm{AgNO_3}$ is sufficient to reach the point of zero charge of AgI. We then have $10^{-10.5}$ M $\mathrm{I}^-$ in solution, i.e. 100.000 times less than $\mathrm{Ag}^+$. The surface charge of AgI can thus be changed by addition of $\mathrm{Ag}^+$ or $\mathrm{I}^-$. The two ions are called the "potential determining ions" of AgI.

How is the surface potential quantitatively related to the concentration of the potential determining ions? In equilibrium, the electrochemical potential of $\mathrm{Ag}^+$ ions at the crystal surface is equal to that in solution:

$$\mu_L^0(Ag^+) + RT \cdot \ln a_L(Ag^+) + F_A\varphi_L = \mu_C^0(Ag^+) + RT \cdot \ln a_C(Ag^+) + F_A\varphi_C \tag{5.12}$$

Here, $\mu_C^0(Ag^+)$and $\mu_L^0(Ag^+)$ are the chemical standard potentials at the crystal surface and in solution, respectively. $\varphi_C$ and $\varphi_L$ are the Galvani potentials in the crystal and in solution. In particular, Eq. (5.12) is valid at the point of zero charge:

$$\mu_L^0(Ag^+) + RT \cdot \ln a_L^{pzc}(Ag^+) = \mu_C^0(Ag^+) + RT \cdot \ln a_C^{pzc}(Ag^+) + F_A \cdot \Delta\chi^{pzc} \tag{5.13}$$

$\Delta\chi^{pzc}$ is the difference of the Galvani potentials, which is caused solely by dipoles in the interface, not by free charges. If we subtract the two equations from each other we get

$$RT \cdot \ln \frac{a_L(Ag^+)}{a_L^{pzc}(Ag^+)} = F_A \cdot (\varphi_C - \varphi_L - \Delta\chi^{pzc}) \tag{5.14}$$

It is assumed that $a_C(Ag^+) = a_C^{pzc}(Ag^+)$. The expression in brackets is the surface potential $\psi_0$. Thus we obtain the Nernst equation:

$$\psi_0 = \frac{RT}{F_A} \cdot \ln \frac{a_L(Ag^+)}{a_L^{pzc}(Ag^+)} \tag{5.15}$$

The concentration of $\mathrm{Ag}^+$ (and thus that of $\mathrm{I}^-$) determines the surface potential. An increase in the concentration by a factor of 10 causes an increase in the surface potential of $\psi_0 = RT/F_A \cdot \ln 10 = 59$ mV at 25 C°.

During the derivation we assumed that $a_C(Ag^+) = a_C^{pzc}(Ag^+)$. That means that during the charging of the AgI surface, the activity of the $\mathrm{Ag}^+$ ions on the surface does not change.

This assumption is justified to a large extent because the total number of $Ag^+$ ions on the surface changes only slightly. The relative number of ions, i.e., the number of additionally adsorbed ions on variation of the potential, is very small.

■ **Example 5.1.** In an uncharged, perfectly crystalline AgI surface, the distance between $Ag^+$ ions is $\approx 0.4$ nm. This corresponds to a surface area for each $Ag^+$ ion of 0.16 $nm^2$. In every $nm^2$ there are 6.25 $Ag^+$ ions. How much does the density of the $Ag^+$ ions increase, if the potential increases by 100 mV? Conditions: 25°C, 1 mM $KNO_3$ background electrolyte.

In order to estimate the surface charge we use the Grahame equation $\sigma = \sqrt{8c_0\varepsilon\varepsilon_0 k_B T} \cdot \sinh(e\psi_0/2k_B T)$. With

$$\frac{e\psi_0}{2k_B T} = \frac{1.60 \times 10^{-19}\,\mathrm{As} \cdot 0.1\,\mathrm{V}}{2 \cdot 1.38 \times 10^{-23}\,\mathrm{JK}^{-1} \cdot 298\,\mathrm{K}} = 1.95 \tag{5.16}$$

and $c_0 = 6.02 \times 10^{20}\,\mathrm{L}^{-1} = 6.02 \times 10^{23}\,\mathrm{m}^{-3}$ we get

$$\begin{aligned} \sigma &= \left(8 \cdot 6.02 \times 10^{23}\mathrm{m}^{-3} \cdot 78.4 \cdot 8.85 \times 10^{-12}\frac{\mathrm{C}^2}{\mathrm{Jm}} \cdot 1.38 \times 10^{-23}\mathrm{J} \cdot 298\right)^{1/2} \\ &\quad \cdot \sinh 1.95 \qquad (5.17) \\ &= 3.71 \times 10^{-3}\frac{\mathrm{C}}{\mathrm{m}^2} \cdot \sinh 1.95 = 0.0128\,\frac{\mathrm{C}}{\mathrm{m}^2} \end{aligned}$$

If we divide this value by unit charge we get an ion density of 0.080 $Ag^+$ ions per $nm^2$. The number of $Ag^+$ ions at the surface changes only by 1.3%.

In the previous section the mercury electrode has been described. If no redox pairs (e.g. $Fe^{2+}$ and $Fe^{3+}$) are in solution and if we exclude gas reactions, the mercury electrode is completely polarizable. Polarizable means: If a potential is applied, a current flows only until the electric double layer has formed. No electrons are transferred from mercury to molecules in the solution and vice versa. The other extreme is a completely reversible electrode, for which the AgI electrode is an example. Each attempt to change the potential of an AgI electrode leads to a current because the equilibrium potential is fixed by the concentrations of $Ag^+$ or $I^-$ according to the Nernst equation.

As an example we consider the electrochemical cell

$$\mathrm{Ag} \mid \mathrm{H_2(Gas)} \mid \mathrm{HNO_3},\ \mathrm{AgNO_3} \mid \mathrm{AgI} \mid \mathrm{Ag}$$

A silver wire is dipped into an aqueous solution containing $HNO_3$ and $AgNO_3$ at equal concentrations. Hydrogen gas is bubbled around the wire so that we have a hydrogen electrode on one side. The other electrode consists of a silver wire coated with AgI. Let us assume a negative potential is applied to the right electrode. At the silver iodide electrode the reactions

$$\mathrm{AgI} + \mathrm{e}^- \rightarrow \mathrm{Ag} + \mathrm{I}^- \quad \text{and} \quad \mathrm{Ag}^+ + \mathrm{e}^- \rightarrow \mathrm{Ag}$$

take place. Iodide ions go into solution, silver precipitates. A Faraday current flows, the concentration of $Ag^+$ at the silver iodide electrode increases, that of $I^-$ decreases until the Nernst equation is satisfied again.

### 5.2.3 Oxides

A third mechanism of surface charging dominates for oxides (e.g. $SiO_2$, $TiO_2$, $Al_2O_3$) [85], proteins, and many water soluble polymers. At the surface of these substances there are groups that can dissociate. They take up or release a proton depending on pH. Examples are hydroxyl, carboxyl, sulfate, and amino groups. The potential determining ions are $OH^-$ and $H^+$.

To calculate the surface potential we consider the simplest example of a surface with one dissociable group. Dissociation leads to a negatively charged group according to

$$\sim \mathrm{AH} \rightleftharpoons \sim \mathrm{A}^- + \mathrm{H}^+ \quad \text{with} \quad K_A = \frac{[A^-]\,[H^+]_{local}}{[AH]} \tag{5.18}$$

The surface concentrations of the negatively charged dissociated groups and the neutral nondissociated groups are given in mol per surface area and not in mol per volume.

$[H^+]_{local}$ is the local proton concentration in the solution directly at the surface. This local concentration can differ from the concentration in the volume phase: if the surface is charged it either attracts protons (by negative surface charge) or it repels protons (by positive surface charge). A relation between the local concentration and the bulk concentration can be found with the help of the Boltzmann factor:

$$\left[H^+\right]_{local} = \left[H^+\right] \cdot e^{-\frac{e\psi_0}{k_B T}} \tag{5.19}$$

Substituting Eq. (5.19) into Eq. (5.18) and taking the logarithm leads to

$$\log K_A = \log \frac{[A^-]}{[AH]} + \log\left[H^+\right] - \frac{e\psi_0 \cdot \log e}{k_B T} \tag{5.20}$$

Since the $pK$ is the negative logarithm of the dissociation constant we get

$$-pK_A = \log \frac{[A^-]}{[AH]} - pH - 0.434 \cdot \frac{e\psi_0}{k_B T} \tag{5.21}$$

or

$$\psi_0 = 2.30 \cdot \frac{RT}{F_A} \cdot (pK_A - pH) + 2.30 \cdot \frac{RT}{F_A} \log \frac{[A^-]}{[AH]} \tag{5.22}$$

At 25°C we have

$$\psi_0 = 59\,\mathrm{mV} \cdot \left[(pK_A - pH) + \log \frac{[A^-]}{[AH]}\right] \tag{5.23}$$

At low pH, many oxides are positively charged. This is at least partly attributable to the process

$$\sim \mathrm{AOH}_2^+ \rightleftharpoons \sim \mathrm{AOH} + \mathrm{H}^+ \tag{5.24}$$

The model was further extended and the binding of other ions at certain binding sites on the solid was allowed [86–90]. These ions partially compete for binding sites with $H^+$ or $OH^-$.

| Substance | pzc |
|---|---|
| $SiO_2$ | 1.8–3.4 |
| $TiO_2$ | 2.9–6.4 |
| $Al_2O_3$ | 8.1–9.7 |
| $MnO_2$ | 1.8–7.3 |
| $Fe_3O_4$ | 6.0–6.9 |
| $\alpha$-$Fe_2O_3$ | 7.2–9.5 |

**Table 5.1:** Points of zero charge for various oxides.

Table 5.1 shows the points of zero charge of several oxides taken from [8]. The point of zero charge is the pH at which the surface charge is zero. In each case a range is given and not a unique value because most oxides occur in different structures and different sorts and concentrations of background electrolytes were used. A comprehensive list of points of zero charge for various oxides can be found in Ref. [91].

As an example, the variation of the surface charge with changing pH is shown in Fig. 5.5 for silicon nitride. Silicon nitride is oxidized at its surface. Charging is primarily caused by the binding and dissociation of $H^+$ according to $\sim SiOH_2^+ \rightleftharpoons \sim SiOH + H^+$ and $\sim SiOH \rightleftharpoons \sim SiO^- + H^+$. For this reason the surface charge becomes increasingly negative with increasing pH [92].

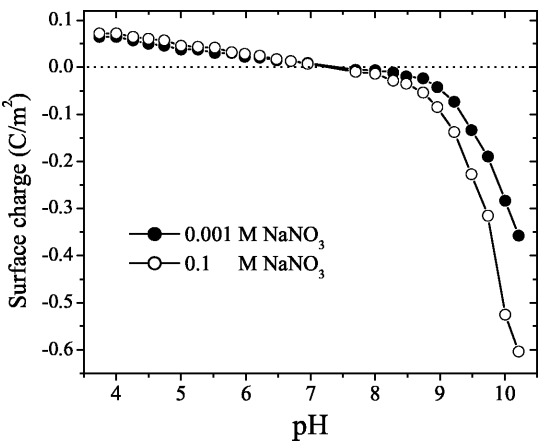



**Figure 5.5:** Surface charge density of commercial silicon nitride in $NaNO_3$ aqueous solutions as a function of pH. Redrawn after Ref. [92].

### 5.2.4 Mica

Clay minerals are formed by two building blocks [93]: tetrahedrons of oxygen with $Si^{4+}$ ions in their centers or octahedrons of oxygen with $Al^{3+}$ or $Mg^{2+}$ in their centers. The tetrahedrons share oxygens and form hexagonal rings. Some oxygen atoms form hydroxyls, in particular when the clay is filled with $Ca^{2+}$. This pattern can be repeated ad infinitum to form flat tetrahedral sheets. Similarly, the octahedrons are linked to form octahedral layers. The tetrahedral and octahedral sheets can be stacked on top of each other in various forms to build the different kinds of clays.

In mica each sheet is composed of three layers (Fig. 5.6). The top and bottom layer are formed by hexagons filled with $Si^{4+}$. The intermediate layer is octahedral and each octahedron is filled with $Al^{3+}$ or $Mg^{2+}$. The sheets are held together by cations. Since this binding is relatively weak, mica can easily be cleaved. An atomically flat surface over huge areas can be obtained. This is the reason why mica is often used as a substrate in microscopy and for surface force experiments.

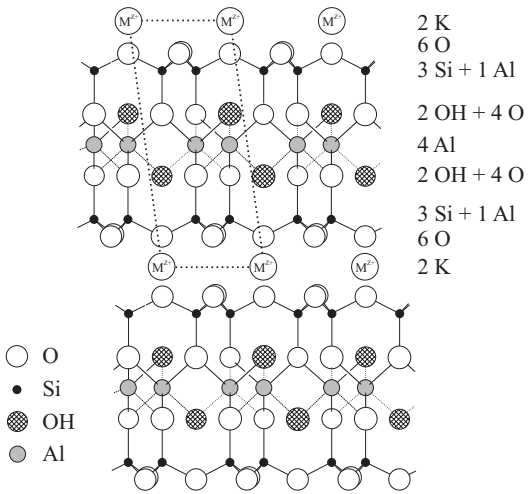



**Figure 5.6:** Molecular structure of muscovite mica ($KAl_2Si_3AlO_{10}(OH)_2$) in side view. A unit cell is indicated by the dashed line. The number of atoms of a certain species in a layer for two unit cells is indicated at the top right. As an average in the third molecular layer from the top one silicon atom is replaced by an aluminium atom.

How does a surface of mica acquire its charge when it is put into a liquid? Here, we first have to distinguish between charging at the edges and charging of the plate. At the edges often hydroxyl groups are responsible for surface charges as on an oxide surface. In water this charge depends sensitively on the pH. Since mica cleaves easily most of the surface will be plate. Charges are acquired by the dissolution of cations. $Si^{4+}$, $Al^{3+}$, and $Mg^{2+}$ dissolve in the liquid leaving a negatively charged mineral behind. Other cations dissolved in the liquid might partially replace the original cations. The total charge is determined by the ion exchange properties of the clay. Charging is only partially due to processes at the surface but more significant is a depletion of positive charge in the mineral. In water this leads to a relatively constant charge, which is not very sensitive to pH [94–97].

### 5.2.5 Semiconductors

Up to this point we have discussed the properties of metals and insulators. Metals have a high conductivity because the electrons are relatively free to move and they can react to an applied potential. Each metal atom contributes one or two electrons so that the density of charge carriers is high. Typically the density of charge carriers is $10^{29}$ per $m^3$. Due to their

high concentration and mobility the electrical conductivity is very high ($10^6$–$10^8$ $\mathrm{Sm}^{-1}$). Insulators, with their tightly bound electrons, are at the other extreme, with conductivities typically lower than $10^{-6}$ $\mathrm{Sm}^{-1}$.

Semiconductors like silicon or germanium are an intermediate case. Their electrons are not as tightly bound as in insulators so that at any given time a small fraction of them will be mobile. In a perfect germanium crystal, for instance at 25°C, about $3 \times 10^{19}$ electrons per $\mathrm{m}^3$ are free. This corresponds to a concentration of $5 \times 10^{-8}$ M or 50 nM. It is much lower than the concentration of charge carriers (cat- and anions) in an aqueous electrolyte solution. Despite this small concentration, the conductivities are of the same order of magnitude, because the electrons in a semiconductor are typically $10^8$ times more mobile than ions in solution.

In a semiconductor we not only have free electrons but also holes (or vacancies left by electrons). The holes are positively charged and move much like the electrons. Therefore the situation is similar to an electrolyte solution with negative and positive charge carriers. The density of electrons can be drastically increased by adding small amounts of P, As, or Sb. The density of holes can be increased by adding B, Al, or Ga. This process is called doping.

How does the electric potential change with respect to distance from the interface? On the solution side the potential decays as described before. We have a Stern and a diffuse layer. For a semiconductor the potential variation at the solid side is also of interest. It also decays and this decay can be described as the decay of the diffuse layer. We only have to replace the salt concentration by the concentration of electrons $c_e$ and holes $c_h$ and we have to use the appropriate dielectric permittivity.

■ **Example 5.2.** Estimate the Debye length of Germanium at 25°C. The dielectric permittivity is $\varepsilon = 16$.

$$\begin{aligned}
\lambda_D &= \sqrt{\frac{\varepsilon\varepsilon_0 k_B T}{2 c_e e^2}} \\
&= \left(\frac{16 \cdot 8.85 \times 10^{-12}\mathrm{AsV}^{-1}\mathrm{m}^{-1} \cdot 4.12 \times 10^{-21}\,\mathrm{J}}{2 \cdot 5 \times 10^{-5}\,\mathrm{molm}^{-3} \cdot 6.02 \times 10^{23}\,\mathrm{mol}^{-1} \cdot (1.60 \times 10^{-19}\mathrm{As})}\right)^{1/2} \\
&= 615\ \mathrm{nm}
\end{aligned}$$

Fig. 5.7 shows schematically the potential across a semiconductor–electrolyte interface. To understand it we have to take two additional effects into account. First, the liquid molecules usually show a preferred orientation at the surface. Their dipole moment causes a jump of the potential. Second, on a solid surface the electrons can occupy surface states. These extra electrons contribute to the potential.

## 5.3 Measuring surface charge densities

### 5.3.1 Potentiometric colloid titration

In most applications we want to know the charge density of dispersed systems. To determine the surface charge of dispersed particles, titration methods can be used. Before we can do a quantitative titration experiment we need to know the specific surface area, that is the total

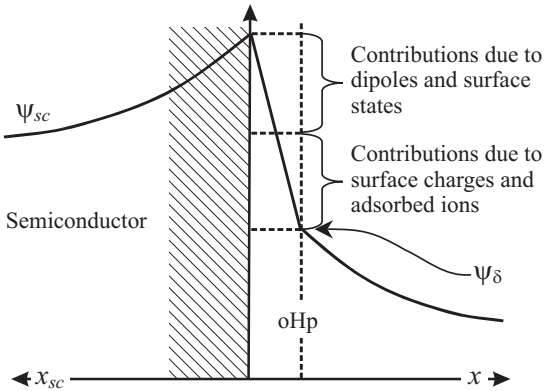



**Figure 5.7:** Potential at a semiconductor–electrolyte interface. The index "sc" refers to semiconductor.

surface area per gram of dispersed material, and the potential-determining ions, either from additional experimental evidence or physicochemical reasoning.

To illustrate the method let us assume that an oxide is titrated in aqueous solution and that $H^+$ and $OH^-$ are the potential-determining ions. A certain amount of the dispersed substance is filled into a cell, which contains a pH electrode. We start the titration at, say, high pH and low ionic strength. The solution could, for instance, contain 1 mM $KNO_3$. An indifferent salt should be used where the ions do not bind specifically to the particles surface. A tiny amount of KOH is added to reach a high pH (Fig. 5.8 point A). Then a small aliquot $HNO_3$ is added. Since $HNO_3$ dissociates completely we know the number of protons added is equal to the amount of $HNO_3$ added. The pH measurement reveals how many protons are in the solution. The difference must be bound to the surface of the oxide. Dividing this number by the total surface area of the particles and multiplying by unit charge, leads to the change in surface charge density (point B). By repeating this procedure many times we reach, step by step, a low pH (C). The pH range is limited by the condition that the amount of $HNO_3$ added should not change the total ionic strength significantly. In addition, we need to make sure that the particles are stable and do not change their properties. This can be controlled by going up with the pH again by adding KOH (to A again). No hysteretic should be observed.

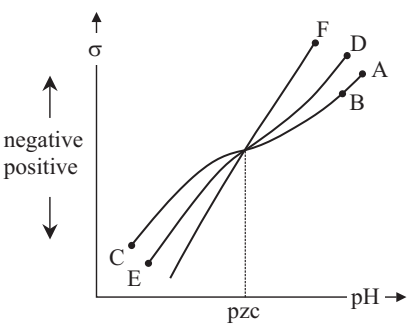



**Figure 5.8:** Schematic result of a potentiometric titration experiment.

Then the salt concentration is increased to, say, 10 mM $KNO_3$ (D). The surface charge increases because the capacitance of the electric double layer increases and so for a given surface potential the charge increases. We repeat the titration with $HNO_3$ (to E) and KOH (back to D). Afterwards the salt concentration is increased to say 0.1 M (F) and the cycle is repeated. The three titration curves intersect at one point — the point of zero charge — if the background ions (in our example $K^+$ and $NO_3^-$) do not bind specifically to the surface [98]. This is essential because, from the titration curves alone, we would only get the shape but not the absolute position of the titration curve. The common intersection gives us the position with respect to the charge axis.

To verify that the salt is really indifferent and no specific binding of the ions occurs, the experiment can be repeated with a different salt. If titration curves measured at the same ionic strength are identical we have good evidence for the absence of specific binding.

In potentiometric colloid titration the amount of potential-determining ions in solution is measured by an appropriate electrode. For oxides, where mainly the pH determines the potential, a glass electrode is a suitable detector. For AgI we could use an AgI electrode etc.

At low ionic strength **conductometric titration** is an alternative approach to determine the amount of ions in solution. It is determined by measuring the electrical conductivity versus the amount of added potential-determining ion. One example is shown in Fig. 5.9. Different amounts of aqueous NaOH solution at a concentration of 5 mM were added to an aquoues dispersion of latex particles. Before adding NaOH, the pH of the dispersion is 3.3 and the conductance is $7.1 \times 10^{-3}\,\Omega^{-1}\mathrm{m}^{-1}$. When adding NaOH the pH first increase gradually, then around an added amount of 2.1 mL the increase becomes steep. Above 3 mL the pH is not affected so much. This indicates that the latex particles have dissociable groups with a pK around 6 on their surface. Correlated with the change in pH is the specific conductance. When adding NaOH one would expect an increase in conductivity because the number of charge carriers increases. This linear increase is indeed observed in the second half of the curve. In the first linearly descending part of the titration curve, protons bind to the surface of the latex particles. As a result, in solution, highly mobile protons are replaced by slower $Na^+$ ions.

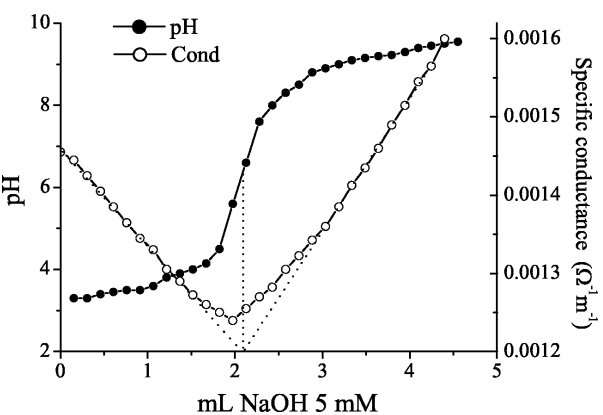



**Figure 5.9:** Conductometric and potentiometric titration of latex particles dispersed in an aqueous medium. Redrawn after Ref. [99].

### 5.3.2 Capacitances

For conducting and completely polarizable surfaces, the capacitance can be measured directly with great precision. In the most simple capacitance measurement, called **chronoamperometry**, a potential step $\Delta U$ is applied to an electrode (Fig. 5.10). A current flows due to charging of the diffuse electrical double layer. The current flows until the capacitance is fully charged. By measuring the current as a function of time and integrating the curve with respect to time we get the charge $Q$. The total capacitance $C$ is easily obtained from $C = C^A\, A = Q/\Delta U$.

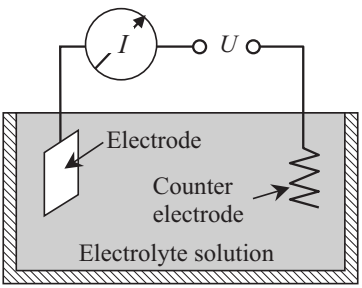



**Figure 5.10:** Schematic experimental set-up to measure capacitance.

■ **Example 5.3.** A potential $\Delta U$ is applied as a step function from 0.1 V to 0.3 V to a platinum wire of 0.5 mm diameter in 0.1 M KCl versus an Ag/AgCl counterelectrode. The counterelectrode has a huge surface area compared to the platinum wire and it is a reversible electrode. For the platinum the applied potential range is too small to allow electrochemical reactions to take place and we have no Faradaic current. The measured current is shown in Fig. 5.11. Integrating the charge (this corresponds to the area underneath the curve) leads to $Q = 59.9$ nC. Dividing by $\Delta U$ we get the a total capacity of $C = 0.299\ \mu$F. The electrode surface area was 0.00196 cm$^2$. This leads to a capacitance per unit area of $C^A = 153\ \mu$F/cm$^2$. The value is higher than expected from Gouy–Chapman theory. One reason is that surface roughness increases the effective surface area. The main reason is, however, that the Gouy–Chapman theory does not adequately describe the capacitance of the double layer, especially at high salt concentrations. A Stern layer has to be taken into account. The example was kindly provided by T. Jenkins.

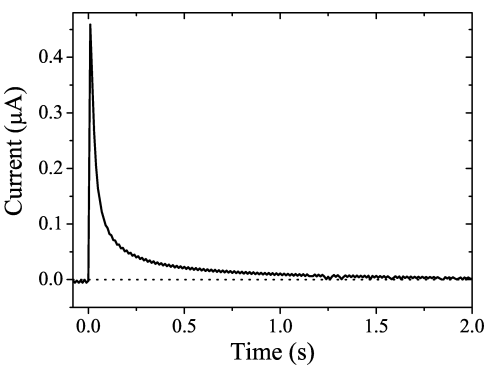



**Figure 5.11:** Chronoamperometric measurement as described in example 5.3.

A standard method to determine capacitance is **cyclic voltammetry**. One electrode, made of the material of interest (with know surface area $A$) and a counter electrode, are introduced into the electrolyte solution. A reference electrode can be used in addition. Then a triangular potential is applied and the electric current is measured. From the current, the capacitance can be calculated.

What can we do for reversible (non polarizable) electrodes? In this case we use the fact that different processes occur at different timescales. Instead of the relatively slow change of the voltage in cyclic voltametry, AC potentials with varying frequencies are applied and the current is detected. The method is called **impedance spectroscopy**. Using impedance spectroscopy, even semiconducting [100] or insulating materials can be analysed by coating them onto metallic electrodes.

## 5.4 Electrokinetic phenomena: The zeta potential

In this section we deal with liquids, which flow along charged solid surfaces. In many cases the surface binds one, two, or several layers of liquid molecules and possibly ions more or less tightly. As a result the shear plane is often not directly at the interface. Only at a distance $\delta$ away from the surface do the molecules start to move. The potential at this distance is called the zeta potential $\zeta$.

The concentration of potential-determining ions at which the zeta potential is zero ($\zeta = 0$) is called the isoelectric point (iep). The isoelectric point is determined by electrokinetic measurements. We have to distinguish it from the point of zero charge (pzc). At the point of zero charge the surface charge is zero. The zeta potential refers to the hydrodynamic interface while the surface charge is defined for the solid–liquid interface.

If a liquid moves tangential to a charged surface, then so-called electrokinetic phenomena arise [101]. Electrokinetic phenomena can be divided into four categories: Electrophoresis, electro-osmosis, streaming potential, and sedimentation potential [102]. In all these phenomena the zeta potential plays a crucial role. The classic theory of electrokinetic effects was proposed by Smoluchowski[2] [103].

### 5.4.1 The Navier–Stokes equation

Electrokinetic phenomena can be understood with the help of two equations: The known Poisson equation and the Navier[3]–Stokes[4] equation. The Navier–Stokes equation describes the movement of a Newtonian liquid, i.e., a liquid whose viscosity does not change when it flows and when it is sheared. In order to make the equation plausible we consider an infinitesimal quantity of the liquid having a volume $dV = dx \cdot dy \cdot dz$ and a mass $dm$. If we want to write Newtons equation of motion for this volume element we have to consider three forces:

- A viscous force, caused by gradients in the shear stress of the fluid $\eta \cdot \mathrm{grad}(\mathrm{div}\vec{v}) \cdot dV = \eta \nabla^2 \vec{v} \cdot dV$.

[2] Marian von Smoluchowski, 1872–1917. Polish physicist, professor in Lemberg and Krakowia.

[3] Claude Louis Marie Henri Navier, 1785–1836. French engineer, professor in Paris.

[4] Sir George Gabriel Stokes, 1819–1903. English mathematician and physicist, professor in Cambridge.

- A possible pressure gradient causes the force $-(\nabla P)\, dV$.
- An electrostatic force, caused by the action of an electric field on the ions in solution $\rho_e \vec{E} dV$.

Here, $P$ is the pressure, $\eta$ is the viscosity, $v$ is the local fluid velocity, $E$ is the electric field strength, and $\rho_e$ is the charge density caused by dissolved ions. According to Newton's law, the sum of these forces is equal to the mass $dm$ times its acceleration:

$$\left(\eta \nabla^2 \vec{v} - \nabla P + \rho_e \vec{E}\right) dV = dm \cdot \frac{d\vec{v}}{dt} \tag{5.25}$$

In many cases a steady-state flow is considered. Then the fluid velocity is constant ($d\vec{v}/dt = 0$) and we get the so-called Navier–Stokes equation:

$$\eta \nabla^2 \vec{v} - \nabla P + \rho_e \vec{E} = 0 \tag{5.26}$$

It is a vector equation. Written in full and in cartesian coordinates it reads:

$$\begin{aligned}
\eta \cdot \left(\frac{\partial^2 v_x}{\partial x^2} + \frac{\partial^2 v_x}{\partial y^2} + \frac{\partial^2 v_x}{\partial z^2}\right) - \frac{\partial P}{\partial x} + \rho_e E_x = 0 \\
\eta \cdot \left(\frac{\partial^2 v_y}{\partial x^2} + \frac{\partial^2 v_y}{\partial y^2} + \frac{\partial^2 v_y}{\partial z^2}\right) - \frac{\partial P}{\partial y} + \rho_e E_y = 0 \\
\eta \cdot \left(\frac{\partial^2 v_z}{\partial x^2} + \frac{\partial^2 v_z}{\partial y^2} + \frac{\partial^2 v_z}{\partial z^2}\right) - \frac{\partial P}{\partial z} + \rho_e E_z = 0
\end{aligned} \tag{5.27}$$

In addition, the equation of continuity applies for incompressible liquids:

$$\nabla \vec{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} = 0 \tag{5.28}$$

These are the basic equations describing the flow of a liquid taking dissolved ions and electric fields into account.

### 5.4.2 Electro-osmosis and streaming potential

Let us start by considering a liquid on a planar, charged surface. If we apply an electric field parallel to the surface the liquid begins to move (Fig. 5.12). This phenomenon is called electro-osmosis. Why does the liquid start to move? The charged surface causes an increase in the concentration of counterions in the liquid close to the surface. This surplus of counterions is moved by the electric field towards the corresponding electrode. The counterions drag the surrounding liquid with them and the liquid starts to flow.

When treating this simple case of electro-osmosis mathematically we immediately realize that the $y$ component of the Navier–Stokes equation disappears. All $y$ derivatives are zero because no quantity can change with $y$ due to the symmetry. We further assume that the liquid flows only parallel to the $x$ coordinate that is parallel to the applied field. Then $v_z = 0$ and $v_y = 0$. As a consequence all derivatives of $v_y$ and $v_z$ are zero. From the equation of

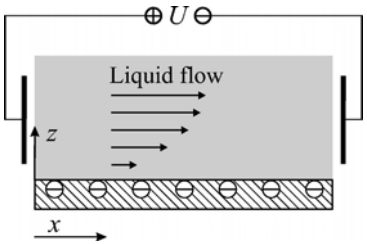



**Figure 5.12:** Electro-osmosis.

continuity we can directly conclude that $\partial v_x/\partial x = 0$. The flow rate does not change with position along the $x$ direction. Since this is valid at all $x$, it follows that $\partial^2 v_x/\partial x^2 = 0$. From the three-component Navier–Stokes equation the only two remaining equations are:

$$\eta \frac{\partial^2 v_x}{\partial z^2} - \frac{\partial P}{\partial x} + \rho_e E_x = 0 \quad \text{and} \quad -\frac{\partial P}{\partial z} + \rho_e E_z = 0 \tag{5.29}$$

The electric field in $x$ direction is applied externally. In $z$ direction the field results from the surface charges. We also assume that no pressure in $x$ direction is applied. Then $\partial P/\partial x$ disappears and after rearranging the first equation we get

$$\rho_e E_x = -\eta \frac{\partial^2 v_x}{\partial z^2} \tag{5.30}$$

Now we use the Poisson equation (4.1), $d^2\psi/dz^2 = -\rho_e/\varepsilon\varepsilon_0$, for the charge density:

$$E_x \varepsilon\varepsilon_0 \frac{d^2\psi}{dz^2} = \eta \frac{d^2 v_x}{dz^2} \tag{5.31}$$

This equation can be integrated twice in $z$. The integration starts at a point far away from the surface, where $\psi = 0$ and $v_x$ has a stationary value $v_0$, up to the shear plane at a distance $\delta$ from the surface where $v_x = 0$ and $\psi = \zeta$. Note that, far from the plane, $d\psi/dz = 0$ and $dv_x/dz = 0$.

$$1^{st}\text{integration}: \quad E_x \varepsilon\varepsilon_0 \cdot \int_{\infty}^{z} \frac{d^2\psi}{dz'^2} dz' = E_x \varepsilon\varepsilon_0 \frac{d\psi}{dz} = \eta \cdot \int_{\infty}^{z} \frac{d^2 v_x}{dz'^2} dz' = \eta \frac{dv_x}{dz} \tag{5.32}$$

$$2^{nd}\text{integration}: \quad E_x \varepsilon\varepsilon_0 \cdot \int_{\infty}^{\delta} \frac{d\psi}{dz'} dz' = E_x \varepsilon\varepsilon_0 \zeta = \eta \cdot \int_{\infty}^{\delta} \frac{dv_x}{dz'} dz' = -\eta v_0 \tag{5.33}$$

It follows that

$$v_0 = -\varepsilon\varepsilon_0 \frac{\zeta E_x}{\eta} \tag{5.34}$$

The flow velocity is proportional to the zeta potential and to the applied field.

We can observe electro-osmosis directly with an optical microscope using liquids, which contain small, yet visible, particles as markers. Most measurements are made in capillaries. An electric field is tangentially applied and the quantity of liquid transported per unit time is measured (Fig. 5.13). Capillaries have typical diameters from 10 $\mu$m up to 1 mm. The diameter is thus much larger than the Debye length. Then the flow rate will change only close to a solid–liquid interface. Some Debye lengths away from the boundary, the flow rate is constant. Neglecting the thickness of the electric double layer, the liquid volume $V$ transported per time is

$$\frac{dV}{dt} = \pi r_C^2 v_0 = \pi r_C^2 \varepsilon\varepsilon_0 \frac{E\varsigma}{\eta} \tag{5.35}$$

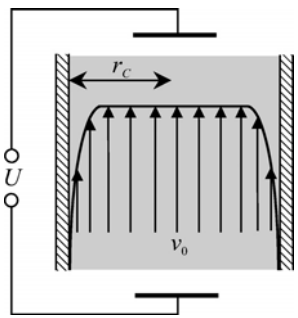



**Figure 5.13:** Electro-osmotic flow profile in a capillary.

Electro-osmosis is widely used in microfluidics to drive aqueous media through thin channels.

■ **Example 5.4.** We apply a potential of 1 V across a capillary of 1 cm length and of 100 $\mu$m diameter. The field should decrease linearly, i.e. $E = 100$ V/m. The capillary is filled with water and the zeta potential is 0.05 V. How much electrolyte flows through the capillary per unit time?

$$\begin{aligned}\frac{dV}{dt} &= \pi \cdot \left(50 \times 10^{-6}\mathrm{m}\right)^2 \cdot 78.5 \cdot 8.85 \times 10^{-12} \frac{\mathrm{C}^2}{\mathrm{Vm}} \cdot \frac{100\,\mathrm{Vm}^{-1} \cdot 0.05\,\mathrm{V}}{0.001\,\mathrm{kgs}^{-1}\mathrm{m}^{-1}} \\ &= 2.73 \times 10^{-14}\,\frac{\mathrm{m}^3}{\mathrm{s}}\end{aligned}$$

With a volume of the capillary of $\pi \cdot \left(50 \times 10^{-6}\mathrm{m}\right)^2 \cdot 0.01\,\mathrm{m} = 7.85 \times 10^{-11}\mathrm{m}^3$ the liquid in the capillary is exchanged in 52 min.

The inverse effect to electro-osmosis is the build-up of a streaming potential. In this case the liquid is pressed through a capillary (more generally along a charged wall). The liquid drags the charges of the electrical double layer with it. As a result some counterions accumulate at the end of the capillary and generate a potential $\Delta U$ — the streaming potential — between the two ends of the capillary. If the radius is substantially larger than the Debye length, it is easily possible to calculate the streaming potential between the beginning and the end of the capillary (Ref. [3] p. 379):

$$\Delta U = \frac{\varepsilon_0 \varepsilon \zeta}{\eta \kappa_e} \cdot \Delta P \tag{5.36}$$

Here, $\Delta P$ is the applied pressure and $\kappa_e$ is the electrical conductivity of the electrolyte (in $\Omega^{-1}\mathrm{m}^{-1}$). Streaming potential measurements are routinely used to characterize the electrical properties of surfaces (e.g. [104]).

### 5.4.3 Electrophoresis and sedimentation potential

In electrophoresis we consider the motion of charged particles in electric fields (in contrast, the movement of ions is treated in electrochemistry under the keyword "ionic conductivity"). Electrophoresis is of great practical importance in biochemistry because it is used to isolate proteins.

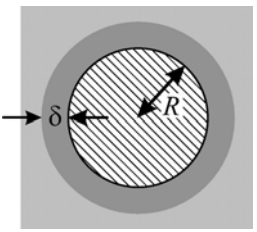



**Figure 5.14:** Bound layer of liquid molecules and ions around a spherical particle.

As an example, we consider the motion of a spherical particle. $\zeta$ is the potential at a distance $\delta$ from the particle surface. The distance $R + \delta$ is the hydrodynamic radius of the particle. It can be larger than the particle radius due to the binding of liquid molecules or ions. Up to this distance the surface charge density amounts to $\sigma_\delta$. The entire charge of the particle is $Q = 4\pi(R+\delta)^2\sigma_\delta \approx 4\pi R^2\sigma_\delta$. An electric field of strength $E$ causes a force $QE$. At constant drift velocity, $v_0$, the force is compensated by the friction force

$$F_F = 6\pi\eta R v_0 \tag{5.37}$$

This is Stokes[5] law of hydrodynamic friction. Equating the two expressions leads to

$$v_0 = \frac{QE}{6\pi\eta R} \tag{5.38}$$

Using the Grahame equation we can express the total charge by the potential. For low potentials $\sigma_\delta = \varepsilon\varepsilon_0\zeta/\lambda_D$ (Eq. 4.27) is valid. Multiplying the charge density by the surface area of the sphere $4\pi R^2$ leads to $Q = 4\pi R^2\varepsilon\varepsilon_0\zeta/\lambda_D$.

This simple equation is, however, only valid for $R \gg \lambda_D$. If the radius is not much larger than the Debye length we can no longer treat the particle surface as an almost planar surface. In fact, we can no longer use the Gouy–Chapman theory but have to apply the theory of Debye and Hückel. Debye and Hückel explicitly considered the electric double layer of a sphere. A result of their theory is that the total surface charge and surface potential are related by

$$Q = 4\pi\varepsilon\varepsilon_0 R\zeta\left(1 + \frac{R}{\lambda_D}\right) \tag{5.39}$$

[5] Sir George Gabriel Stokes, 1819–1903. British mathematician and physicist, professor in Cambridge.

With the help of this relationship we can replace the charge with the zeta potential:

$$v_0 = \frac{2\varepsilon\varepsilon_0\zeta E}{3\eta} \cdot \left(1 + \frac{R}{\lambda_D}\right) \tag{5.40}$$

Until now we have ignored an important factor. The electric field affects not only the surface charges of the particle, but also the ions in the electrical double layer. The counterions in the double layer move in a direction opposite to the motion of the particle. The liquid transported by them inhibits the particle motion. This effect is called *electrophoretic retardation*. Therefore the equation is only valid for $\lambda_D \ll R$. Generally the term in brackets must be replaced by a function $f(R/\lambda_D)$ and we get [105, 106]

$$v_0 = \frac{2\varepsilon\varepsilon_0\zeta E}{3\eta} \cdot f\left(\frac{R}{\lambda_D}\right) \tag{5.41}$$

The "opposite" effect to electrophoresis is the generation of a sedimentation potential. If a charged particle moves in the gravitational field or in a centrifuge, an electric potential arises — the sedimentation potential. While the particle moves, the ions in the electric double layer lag somewhat behind due to the liquid flow. A dipole moment is generated. The sum of all dipoles causes the sedimentation potential.

## 5.5 Types of potentials

At this point it might help to define precisely what kind of electrical potentials occur. Electrical potential differences between two points, say A and B, are defined as the electrical work done in transporting a unit charge from A to B; the potential is the work divided by the unit charge. To each thermodynamically stable phase we can assign an internal or **Galvani potential** $\varphi$. The internal potential is defined by the work, which is required to bring a test charge from an infinite distance and from a vacuum into the interior of the phase far from the phase boundary.

The transport of the test charge can be pictured as occurring in two steps (Fig. 5.15): First, the charge is brought close (approx. 1 $\mu$m) to the interface. This step is related to the **Volta** or **external potential** $\psi$. It depends on the exact position. Also in the case of the electric double layer we used the symbol for the Volta potential. In this case the reference point is not in vacuum but in the bulk liquid phase far away from the interface $\psi\,(x = \infty) = 0$.

In the second step the charge arrives at the internal phase passing through the interface. The associated potential is known as the **surface potential jump** $\chi$ (also called surface potential, surface electrical potential, etc.). It is determined by dipoles aligned at the interface and by surface charges. It is not identical with the Volta potential difference (also sometimes called the surface potential) that has so far been used for the description of the electrical double layer. For the treatment of the electrical double layer, dipoles did not play a role. In particular in water, however, the aligned water molecules contribute substantially to the surface potential jump $\chi$. The Galvani potential, Volta potential, and surface potential jump are related by $\varphi = \psi + \chi$.

The Volta potentials of individual phases and Volta potential differences $\Delta\psi$ between two phases can be measured. In contrast, the Galvani potential and accordingly the surface potential of a single phase cannot be measured. Explanation: If we bring a test charge (an ion or an

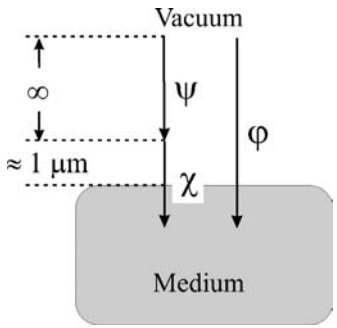



**Figure 5.15:** Illustration of the Galvani $\varphi$, Volta $\psi$, and surface potential $\chi$.

electron) into a medium there is always some chemical work in addition to the electrical work. It can be due to van der Waals forces or image charge effects. There is no way of distinguishing the chemical from the electrical work as long as the chemical environment is different in the two phases. For the same reason differences in the Galvani potential $\Delta\varphi$ between two materials and differences in the surface potential $\Delta\chi$ cannot be measured. Experimentally accessible are only $\Delta(\Delta\varphi)$ and $\Delta(\Delta\chi)$ changes of differences in Galvani or surface potentials. An example are electrocapillary measurements.

The **electrochemical potential** $\mu_i^\alpha$ is measurable. It is defined as the work required to bring a particle of the sort $i$ from vacuum into the phase $\alpha$. Experimentally it is, however, impossible to separate the chemical potential from the electrostatic part.

In this context it is instructive to mention the **work function** $\Phi$. Here, $e\Phi$ indicates the work that is necessary to transport an electron from the Fermi level of a solid into vacuum far away from the surface [107]. The work function can be measured for example by ultraviolet photon spectroscopy (for a discussion see Ref. [108]).

As an example, let us consider the potential in an electrochemical cell. For the sake of simplicity an electrode $A'$ is made of the same metal as the connection to the second electrode. The measured potential corresponds to the difference in the Galvani potentials

$$U = \varphi^{A'} - \varphi^{A} \equiv {}^{A'}\Delta^{A}\varphi \tag{5.42}$$

The last equality is only to introduce the notation used from now on. One could object that now differences of the Galvani potential can be measured, although it has been stated that this is impossible. In reality, the Galvani potential difference between material $A$ and material $A'$ is not measured (it is zero), but only the change in the potential difference due to the presence of the other materials.

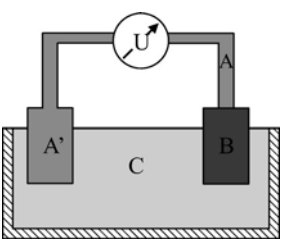



**Figure 5.16:** Electrochemical cell.

The sum of the voltages over all transitions must be zero, in accordance with Kirchhoffs law. From this we get

$$U = {}^{A'}\Delta^{A}\Phi = {}^{A}\Delta^{B}\varphi + {}^{B}\Delta^{C}\varphi + {}^{C}\Delta^{A'}\varphi \tag{5.43}$$

The voltage can also be expressed by the electrochemical potentials of the electrons:

$$U = \frac{\mu_e^{A'} - \mu_e^{A}}{e} = \frac{\mu_e^{A'} - \mu_e^{B}}{e} \tag{5.44}$$

The last equation is valid since metal A and metal B are in contact and the electrons can flow freely over the border. Therefore $\mu_e^A = \mu_e^B$.

## 5.6 Summary

- Valuable information about the properties of electrical double layers can be obtained from electrocapillary experiments. In an electrocapillary experiment the surface tension of a metal surface versus the electrical potential is measured. The capacitance and the point of zero charge are obtained. Surface charge densities for disperse systems can be determined by potentiometric and conductometric titration.
- Different processes are responsible for the build-up of surface charge in different materials such as metals, weakly soluble salts, oxides, and clay minerals.
- The zeta potential is the potential of a solid surface at the shear plane of the surrounding liquid. The zeta potential is relevant in electrokinetic phenomena. In electro-osmosis a liquid flow is induced by applying an electrical potential parallel to a charged surface. Streaming potentials are caused by a liquid flow tangential to a charged surface. Electrophoresis is the movement of a charged particle in an electric field.

## 5.7 Exercises

1. What is the Gibbs free energy of the electrostatic double layer around a sphere of radius $R$ which has a surface charge density $\sigma$? Derive an equation where the Gibbs free energy of one sphere is given as a function of the total charge. You can assume a low surface potential.
2. For a microfluidic application, a capillary of 10 $\mu$m radius and 5 cm length was fabricated in glass. The zeta potential of this glass in 0.01 M KCl aqueous solution at neutral pH is $-30$ mV. A potential of 5 V is applied along the capillary. How fast and in which direction does the liquid flow?
3. To observe the flow, small spherical polystyrene particles of 50 nm radius which are fluorescently labeled, are added. To keep them dispersed they have sulfate groups on their surface. This leads to a zeta potential of $-20$ mV. How fast and in which direction do these particles move? A good marker should move with the same speed as the liquid flow. Was it a good idea to use these particles as markers?

# 6 Surface forces

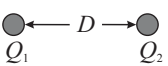
A whole range of phenomena in interface science revolve around the effect of surface forces. Many practical applications in colloid science come down to the problem of controlling the force between colloidal particles, between particles and surfaces, and between two surfaces. For this reason scientists have devoted considerable effort to understanding surface forces and being able to influence them.

When talking about "surface forces" the first association is that of two solid particles interacting in a fluid medium. Practically this is of direct relevance when dealing with the stability of sols. If attractive interactions dominate, the particles aggregate, if repulsive forces dominate, the dispersion is stable. The subject "surface forces" is, however, more general. It, for example, includes the interaction between two liquid–gas interfaces of a liquid lamella which is relevant for the stability of foams. Here, the two opposing liquid–vapor interfaces interact. If they are attractive, the liquid lamella is unstable and ruptures. For a repulsive interaction the lamella is stable. Another example are thin liquid films on solid surfaces. Here, the solid–liquid and the liquid–vapor interfaces interact. A repulsive interaction leads to a stable film which tends to wet the solid surface while attractive interactions destabilize thin films. This will also be discussed in Chapter 7. A very good introduction about surface forces is Ref. [109]. A brief historical overview is given in Ref. [110].

## 6.1 Van der Waals forces between molecules

Forces between macroscopic objects result from a complex interplay of the interaction between molecules in the two objects and the medium separating them. The basis for an understanding of intermolecular forces is the Coulomb[1] force. The Coulomb force is the electrostatic force between two charges $Q_1$ and $Q_2$:

$$F = \frac{Q_1 Q_2}{4\pi\varepsilon\varepsilon_0 D^2} \tag{6.1}$$




The potential energy between two electrical charges which are a distance $D$ apart is

$$W = \frac{Q_1 Q_2}{4\pi\varepsilon\varepsilon_0 D} \tag{6.2}$$

For charges with opposite sign the potential energy is negative. They reduce their energy when they get closer. If the two charges are in a medium, the dielectric permittivity $\varepsilon$ is higher than one and the electrostatic force is reduced accordingly.

[1] Charles Augustin Coulomb, 1736–1806. French physicist and engineer.

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

■ **Example 6.1.** The potential energy between $Na^+$ and $Cl^-$, being 1 nm apart, in a vacuum is

$$W = -\frac{\left(1.60 \times 10^{-19}\mathrm{C}\right)^2}{4\pi \cdot 8.85 \times 10^{-12}\mathrm{AsV^{-1}m^{-1}} \cdot 10^{-9}\mathrm{m}} = -2.30 \times 10^{-19}\ \mathrm{J}$$

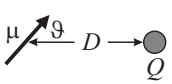
This is 56 times higher than the thermal energy $k_B T = 4.12 \times 10^{-21}$ J at room temperature.

Most molecules are not charged. Still, the electric charge is often not distributed evenly. A molecule can have a more negative side and a more positive side. In carbon monoxide, for example, the oxygen is more negative than the carbon atom. To first order, the electric properties of such molecules are described by the so-called dipole moment. For the most simple case of two opposite charges $Q$ and $-Q$ being a distance $d$ apart, the dipole moment $\mu$ is given by $\mu = Q \cdot d$. It is given in units of Cm. Often the old unit "Debye" is used. 1 Debye is equal to a positive and a negative unit charge being 0.21 Å apart; it is $1\ \mathrm{D} = 3.336 \times 10^{-30}$ Cm. The dipole moment is a vector which points from minus to plus. If we have more than two charges we have to integrate the charge density $\rho_e$ over the whole volume of the molecule, which leads to the general definition of the dipole moment:

$$\vec{\mu} = \int \rho_e(\vec{r}) \cdot \vec{r} \cdot dV \tag{6.3}$$

Let us now return to intermolecular interactions. If more than two charges are present the net potential energy of a charge can be calculated by summing up the contributions of all the other charges. This is called the superposition principle. Using this superposition principle we can calculate the potential energy between a dipole and a single charge:

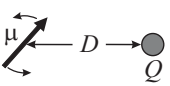
$$W = -\frac{Q\mu \cdot \cos\vartheta}{4\pi\varepsilon_0 D^2} \tag{6.4}$$




Here, we have assumed that the distance $D$ is large compared to the extension of the dipole.

In practice a molecule with a dipole moment is often mobile. If the dipole is free to rotate and close to a positive charge it tends to rotate until its negative pole points towards the positive charge. On the other hand, thermal fluctuations drive it away from a perfect orientation. On average, a net preferential orientation remains and the dipole is attracted by the monopole. The average potential energy is

$$W = -\frac{Q^2\mu^2}{6(4\pi\varepsilon_0)^2 k_B T D^4} \tag{6.5}$$



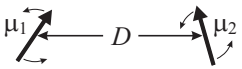

■ **Example 6.2.** Calculate the potential energy between $Na^+$ and a water molecule (dipole moment $6.17 \times 10^{-30}$ Cm) being 1 nm apart in vacuum at 25°C.

$$\begin{aligned} W &= -\frac{\left(1.60 \times 10^{-19}\mathrm{C}\right)^2 \left(6.17 \times 10^{-30}\mathrm{Cm}\right)^2}{6\left(4\pi \cdot 8.85 \times 10^{-12}\mathrm{AsV^{-1}m^{-1}}\right)^2 \cdot 4.12 \times 10^{-21}\mathrm{J} \cdot \left(10^{-9}\mathrm{m}\right)^4} \\ &= -3.20 \times 10^{-21}\ \mathrm{J} \end{aligned}$$

This is roughly equal to the thermal energy $k_B T$.

Two freely rotating dipoles attract each other because they preferentially orient with their opposite charges facing each other. This randomly oriented dipole–dipole interaction is often referred to as the Keesom energy[2]:

$$W = -\frac{C_{orient}}{D^6} = -\frac{\mu_1^2\mu_2^2}{3(4\pi\varepsilon_0)^2 k_B T D^6} \tag{6.6}$$




The coefficient $C_{orient}$ is independent of the distance. For two water molecules which are 1 nm apart, the Keesom energy is, for example, $-9.5 \times 10^{-24}$ J.

All expressions reported so far give the Helmholtz free energies of interaction because they were derived under constant volume conditions. Until now free energy and internal energy were identical. For the randomly oriented dipole–dipole interaction, entropic effects, namely the ordering of one dipole by the field of the other dipole, contribute to the free energy. If one dipole approaches another, one-half of the internal energy is taken up in decreasing the rotational freedom of the dipoles as they become progressively more aligned. For this reason, the free energy as given in Eq. (6.6) is only half of the internal energy. The Gibbs free energy is in this case equal to the Helmholtz free energy.

When a charge approaches a molecule without a static dipole moment, all energies considered so far would be zero. Nevertheless, there is an attractive force. Reason: The monopole induces a charge shift in the non-polar molecule. An induced dipole moment arises, which interacts with the charge. The Helmholtz free energy is

$$W = -\frac{Q^2\alpha}{2(4\pi\varepsilon_0)^2 D^4} \tag{6.7}$$

Here, $\alpha$ is the polarizability in $C^2m^2J^{-1}$. The polarizability is defined by $\mu_{ind} = \alpha E$, where $E$ is the electric field strength . Often it is given as $\alpha/4\pi\varepsilon_0$ in units of $Å^{-3}$.

■ **Example 6.3.** The polarizability of a water molecule is $1.65\times10^{-40}$ $C^2m^2J^{-1}$. Which dipole moment is induced by a unit charge which is 1 nm away and what is the potential energy between the two? The electric field of a point charge at a distance $D$ is

$$E = \frac{Q}{4\pi\varepsilon_0 D^2} = 1.44 \times 10^9 \,\frac{\mathrm{V}}{\mathrm{m}}$$

The induced dipole moment is

$$\mu_{ind} = 1.65 \times 10^{-40}\,\mathrm{C^2J^{-1}m^2} \cdot 1.44 \times 10^9\,\mathrm{Vm^{-1}} = 2.38 \times 10^{-31}\,\mathrm{Cm}$$

and the potential energy is $-1.71\times10^{-22}$ J.

Also a molecule with a static dipole moment interacts with a polarisable molecule. If the dipole can freely rotate, the Helmholtz free energy is

$$W = -\frac{C_{ind}}{D^6} = -\frac{\mu^2\alpha}{(4\pi\varepsilon_0)^2 D^6} \tag{6.8}$$

[2] Wilhelmus Hendrik Keesom, 1876–1956. Dutch physicist, professor in Utrecht and Leiden.

for different molecules. If the molecules are identical, a factor of two has to be inserted. This randomly oriented induced dipole interaction is called the Debye interaction.

All energies considered so far can be calculated using classical physics. They fail to explain the attraction between nonpolar molecules. That such an attraction exists is evident because all gases condense at some temperature. Responsible for this attraction is the so called London[3] or dispersion force. To calculate the dispersion force, quantum mechanical perturbation theory is required. An impression about the origin of dispersion forces can be obtained by considering an atom with its positively charged nucleus around which electrons circulate with a high frequency of typically $10^{15}$–$10^{16}$ Hz. At every instant, the atom is therefore polar. Only the direction of the polarity changes with this high frequency. When two such oscillators approach, they start to influence each other. Attractive orientations have higher probabilities than repulsive ones. As an average this leads to an attractive force.

The Helmholtz free energy between two molecules with ionisation energies $h\nu_1$and $h\nu_2$ can be approximated by

$$W = -\frac{C_{disp}}{D^6} = -\frac{3}{2} \cdot \frac{\alpha_1 \alpha_2}{(4\pi\varepsilon_0)^2 D^6} \cdot \frac{h\nu_1\nu_2}{(\nu_1 + \nu_2)} \tag{6.9}$$

Dispersion interactions increase with the polarizability of the two molecules. The optical properties enter in the form of the excitation frequencies. Expression (6.9) only considers one term of a series over dipole transition moments. Usually, however, this term is by far the most dominant one.

Keesom, Debye, and London contributed much to our understanding of forces between molecules [111–113]. For this reason the three dipole interactions are named after them. The van der Waals[4] force is the Keesom plus the Debye plus the London dispersion interaction, thus, all the terms which consider dipole–dipole interactions: $C_{total} = C_{orient} + C_{ind} + C_{disp}$. All three terms contain the same distance dependency: the potential energy decreases with $1/D^6$. Usually the London dispersion term is dominating. Please note that polar molecules not only interact via the Debye and Keesom force, but dispersion forces are also present. In Table 6.1 the contributions of the individual terms for some gases are listed.

So far we have implicitly assumed that the molecules stay so close to each other that the propagation of the electric field is instantaneous. For the London dispersion interaction this is not necessarily true. To illustrate this let us have a closer look at what happens when two molecules interact. In one molecule a spontaneous random dipole moment arises, which generates an electric field. The electric field expands with the speed of light. It polarises the second molecule, whose dipole moment in turn causes an electric field that reaches the first molecule with the speed of light. The process takes place as calculated only if the electric field has enough time to cover the distance $D$ between the molecules before the dipole moment has completely changed again. This takes a time $\Delta t = D/c$, where $c$ is the speed of light. If the first dipole changes faster than $\Delta t$, the interaction becomes weaker. The time during which the dipole moment changes is in the order of $1/\nu$. Hence, only if

$$\frac{D}{c} < \frac{1}{\nu} \tag{6.10}$$

[3] Fritz London, 1900–1954. American physicist of German origin, professor in Durham.

[4] Johannes Diderik van der Waals, 1837–1923. Dutch physicist, professor in Amsterdam.

**Table 6.1:** Contributions of the Keesom, Debye, and London potential energy to the total van der Waals interaction between similar molecules as calculated with Eqs. (6.6), (6.8), and (6.9) using $C_{total} = C_{orient} + C_{ind} + C_{disp}$. They are given in units of $10^{-79}$ Jm$^6$. For comparison, the van der Waals coefficient $C_{exp}$ as derived from the van der Waals equation of state for a gas $\left(P + a/V_m^2\right) \cdot (V_m - b) = RT$ is tabulated. From the experimentally determined constants $a$ and $b$ the van der Waals coefficient can be calculated with $C_{\mathrm{exp}} = 9ab/\left(4\pi^2 N_A^3\right)$ [109] assuming that at very short range the molecules behave like hard core particles. Dipole moments $\mu$, polarizabilities $\alpha$, and the ionization energies $h\nu$ of isolated molecules are also listed.

| | $\mu$ (D) | $\alpha/4\pi\varepsilon_0$ ($10^{-30}$m$^3$) | $h\nu$ (eV) | $C_{orient}$ | $C_{ind}$ | $C_{disp}$ | $C_{total}$ | $C_{exp}$ |
|---|---|---|---|---|---|---|---|---|
| He | 0 | 0.2 | 24.6 | 0 | 0 | 1.2 | 1.2 | 0.86 |
| Ne | 0 | 0.40 | 21.6 | 0 | 0 | 4.1 | 4.1 | 3.6 |
| Ar | 0 | 1.64 | 15.8 | 0 | 0 | 50.9 | 50.9 | 45.3 |
| $CH_4$ | 0 | 2.59 | 12.5 | 0 | 0 | 101.1 | 101.1 | 103.3 |
| HCl | 1.04 | 2.7 | 12.8 | 9.5 | 5.8 | 111.7 | 127.0 | 156.8 |
| HBr | 0.79 | 3.61 | 11.7 | 3.2 | 4.5 | 182.6 | 190.2 | 207.4 |
| HI | 0.45 | 5.4 | 10.4 | 0.3 | 2.2 | 364.0 | 366.5 | 349.2 |
| $CHCl_3$ | 1.04 | 8.8 | 11.4 | 9.5 | 19.0 | 1058 | 1086 | 1632 |
| $CH_3OH$ | 1.69 | 3.2 | 10.9 | 66.2 | 18.3 | 133.5 | 217.9 | 651:0 |
| $NH_3$ | 1.46 | 2.3 | 10.2 | 36.9 | 9.8 | 64.6 | 111.2 | 163.7 |
| $H_2O$ | 1.85 | 1.46 | 12.6 | 95.8 | 10.0 | 32.3 | 138.2 | 176.2 |
| CO | 0.11 | 1.95 | 14.0 | 0.0012 | 0.047 | 64.0 | 64.1 | 60.7 |
| $CO_2$ | 0 | 2.91 | 13.8 | 0 | 0 | 140.1 | 140.1 | 163.6 |
| $N_2$ | 0 | 1.74 | 15.6 | 0 | 0 | 56.7 | 56.7 | 55.3 |
| $O_2$ | 0 | 1.58 | 12.1 | 0 | 0 | 36.2 | 36.2 | 46.0 |

the interaction takes place as considered. The relevant frequencies are those corresponding to the ionization of the molecule, which are typically $3\times10^{15}$ Hz. Thus, for $D > c/\nu \approx 3 \times 10^8/3 \times 10^{15}\mathrm{m} = 10$ nm the van der Waals energy decreases more steeply (i.e. for molecules with $1/D^7$) as for smaller distances. This effect is known as retardation.

## 6.2 The van der Waals force between macroscopic solids

We now move from the interaction between two molecules to the interaction between two macroscopic solids. There are two approaches to calculate the van der Waals force between extended solids: the microscopic and the macroscopic [114].

### 6.2.1 Microscopic approach

The potential energy of the interaction between molecule A and molecule B is

$$W_{AB}(D) = -\frac{C_{AB}}{D^6} \tag{6.11}$$

The minus sign arises because it is an attractive force. $C_{AB}$ is equal to $C_{total}$ and sums up contributions of all three dipole–dipole interactions .

In order to determine the interaction between macroscopic solids, in the first step we calculate the van der Waals energy between a molecule A and an infinitely extended body with a planar surface made of molecules B. This is also of direct relevance in understanding the adsorption of gas molecules to surfaces. We sum up the van der Waals energy between molecule A and all molecules in the solid B. Practically this is done via an integration of the molecular density $\rho_B$ over the entire volume of the solid:

$$W_{Mol/plane} = -C_{AB} \cdot \iiint \frac{\rho_B}{D'^6} dV = -C_{AB}\rho_B \cdot \int_0^\infty \int_0^\infty \frac{2\pi r dr dx}{\left((D+x)^2 + r^2\right)^3} \quad (6.12)$$

We have used cylindrical coordinates (see Fig. 6.1) and we have assumed that the density of molecules B in the solid is constant. With $2rdr = d(r^2)$ we get

$$\begin{aligned} W_{Mol/plane} &= -\pi\rho_B C_{AB} \cdot \int_0^\infty \int_0^\infty \frac{d(r^2)}{\left((D+x)^2 + r^2\right)^3} dx \\ &= -\pi\rho_B C_{AB} \cdot \int_0^\infty \left[ -\frac{1}{2\left((D+x)^2 + r^2\right)^2} \right]_0^\infty dx \\ &= -\frac{\pi\rho_B C_{AB}}{2} \cdot \int_0^\infty \frac{1}{(D+x)^4} dx \\ &= -\frac{\pi\rho_B C_{AB}}{2} \cdot \left[ -\frac{1}{3(D+x)^3} \right]_0^\infty \\ &= -\frac{\pi\rho_B C_{AB}}{6D^3} \end{aligned} \quad (6.13)$$

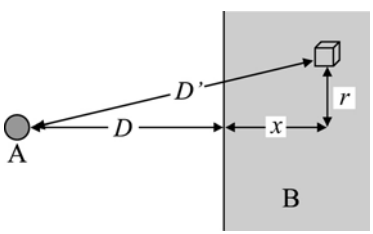



**Figure 6.1:** Calculating the van der Waals force between a macroscopic body and a molecule.

Here we already see a remarkable property of the van der Waals forces: The energy of a molecule and a macroscopic body decreases less steeply than the energy between two molecules. Instead of the $D^{-6}$ dependence the energy decreases proportional to $D^{-3}$.

In the second step we calculate the van der Waals energy between two infinitely extended solids which are separated by a parallel gap of thickness $D$. Therefore, we use Eq. (6.13) and

integrate over all molecules in the solid A:

$$W = -\frac{\pi C_{AB}\rho_B}{6}\iiint \frac{\rho_A}{(D+x)^3}dV = -\frac{\pi C_{AB}\rho_B}{6}\int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty \frac{\rho_A dzdydx}{(D+x)^3} \tag{6.14}$$

Here, $y$ and $z$ are the coordinates parallel to the gap. The integral is infinite because the solids are infinitely large. We have to divide by the area. The van der Waals energy per unit area is

$$\begin{aligned} w &= \frac{W}{A} = -\frac{\pi\rho_A\rho_B C_{AB}}{6}\int_0^\infty \frac{dx}{(D+x)^3} \\ &= -\frac{\pi\rho_A\rho_B C_{AB}}{6}\left[-\frac{1}{2(D+x)^2}\right]_0^\infty \\ &= -\frac{\pi\rho_A\rho_B C_{AB}}{12D^2} \end{aligned} \tag{6.15}$$

With the definition of the so-called Hamaker constant

$$A_H = \pi^2 C_{AB}\rho_A\rho_B \tag{6.16}$$

we get

$$w = -\frac{A_H}{12\pi D^2} \tag{6.17}$$

The force per unit area is equal to the negative derivative of $w$ versus distance:

$$f = -\frac{A_H}{6\pi D^3} \tag{6.18}$$

In the same way it is possible to calculate the van der Waals energy between solids having different geometries. One important case is the interaction between two spheres with radii $R_1$ and $R_2$. The van der Waals energy is [115]

$$W = -\frac{A_H}{6}\cdot\left[\frac{2R_1R_2}{d^2-(R_1+R_2)^2} + \frac{2R_1R_2}{d^2-(R_1-R_2)^2} + \ln\left(\frac{d^2-(R_1+R_2)^2}{d^2-(R_1-R_2)^2}\right)\right] \tag{6.19}$$

where $d$ is the distance between the centers of the spheres. The distance between the surfaces is $D = d - R_1 - R_2$.

In Eq. (6.19) only attractive van der Waals forces were taken into account. At very short range the electron orbitals start to overlap and the molecules repel each other. Henderson considered this short-range repulsion and reports a modified version of Eq. (6.19) [116].

If the radii of the spheres are substantially larger than the distance ($D \ll R_1, R_2$) Eq. (6.19) can be simplified (Ref. [3], p. 543):

$$W = -\frac{A_H}{6D}\cdot\frac{R_1R_2}{R_1+R_2} \tag{6.20}$$

The van der Waals force is the negative derivative:

$$F = -\frac{A_H}{6D^2} \cdot \frac{R_1 R_2}{R_1 + R_2} \tag{6.21}$$

For a sphere and a flat, planar surface the energy and force can be obtained by letting $R_2$ go to infinity:

$$W = -\frac{A_H R}{6D} \qquad \text{and} \qquad F = -\frac{A_H R}{6D^2} \tag{6.22}$$

■ **Example 6.4.** A spherical quartz ($SiO_2$) particle hangs on a planar quartz surface caused by the van der Waals attraction of $F = A_H R/6D^2$. The van der Waals attraction increases linearly with the radius of the sphere. The gravitational force, which pulls the sphere down, is $4\pi R^3 \rho g/3$. It increases cubically with the radius. As a consequence, the behavior of small spheres is dominated by van der Waals forces, while for large spheres gravity is more important. At which radius is the gravitational force so strong that the sphere detaches?

The Hamaker constant is $A_H \approx 6 \times 10^{-20}$ J, the density is $\rho = 3000\,\mathrm{kgm}^{-3}$. For two solids in contact we take the distance to be 1.7 Å, which corresponds to a typical interatomic spacing. We use the condition that, just before the sphere falls down, the van der Waals force is equal to the gravitational force:

$$\frac{A_H R}{6D^2} = \frac{4}{3}\pi R^3 \rho g \;\Rightarrow\; R = \sqrt{\frac{A_H}{8\pi D^2 \rho g}} \;\Rightarrow$$

$$R = \left( \frac{6 \times 10^{-20}\,\mathrm{J}}{8\pi \cdot (1.7 \times 10^{-10}\mathrm{m})^2 \cdot 3000\,\mathrm{kgm}^{-3} \cdot 9.81\,\mathrm{ms}^{-1}} \right)^{1/2} = 1.7\ \mathrm{mm}$$

In reality the value is certainly lower due to surface roughness and contamination on the surfaces. Roughness and contamination layers increase the effective contact distance.

■ **Example 6.5.** Geckos are able to run up and down walls, and even walk along the ceiling head downwards. Therefore they have evolved one of the most versatile and effective mechanisms to adhere to a surface. Geckos accomplish this by using van der Waals forces [117]. Their feet are covered with millions of tiny foot-hairs, called setae; typically they have 14,000 setae per $\mathrm{mm}^2$. Each setae consists of many spatulae (Fig. 6.2) which even out the surface roughness and in this way achieve a high total contact area.

### 6.2.2 Macroscopic calculation — Lifshitz theory

In the microscopic calculation pairwise additivity was assumed. We ignored the influence of neighboring molecules on the interaction between any pair of molecules. In reality the van der Waals force between two molecules is changed by the presence of a third molecule. For example, the polarizability can change. This problem of additivity is completely avoided in the macroscopic theory developed by Lifshitz [118, 119]. Lifshitz neglects the discrete atomic structure and the solids are treated as continuous materials with bulk properties such as the

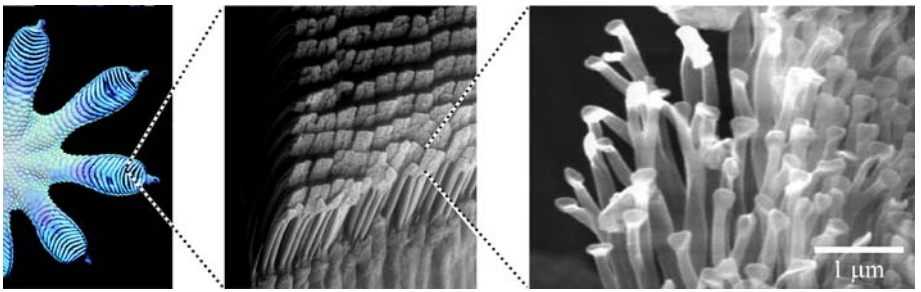



**Figure 6.2:** Picture of a gecko foot (left) and scanning electron microscope images of different resolution of setae (middle) and spatulae (right). Images were kindly provided by K. Autumn.

dielectric permittivity and the refractive index. Fortunately, the same expressions are obtained, in particular, the distance dependencies turn out to be correct. Only the Hamaker constant is calculated in a different way.

Though the original work is difficult to understand very good reviews about the van der Waals interaction between macroscopic bodies have appeared [114, 120]. In the macroscopic treatment the molecular polarizability and the ionization frequency are replaced by the static and frequency dependent dielectric permittivity. The Hamaker constant turns out to be the sum over many frequencies. The sum can be converted into an integral. For a material 1 interacting with material 2 across a medium 3, the non-retarded[5] Hamaker constant is

$$\begin{aligned} A_H &= \frac{3}{4} k_B T \cdot \left( \frac{\varepsilon_1 - \varepsilon_3}{\varepsilon_1 + \varepsilon_3} \right) \cdot \left( \frac{\varepsilon_2 - \varepsilon_3}{\varepsilon_2 + \varepsilon_3} \right) \\ &+ \frac{3h}{4\pi} \cdot \int_{\nu_1}^{\infty} \left( \frac{\varepsilon_1(i\nu) - \varepsilon_3(i\nu)}{\varepsilon_1(i\nu) + \varepsilon_3(i\nu)} \right) \cdot \left( \frac{\varepsilon_2(i\nu) - \varepsilon_3(i\nu)}{\varepsilon_2(i\nu) + \varepsilon_3(i\nu)} \right) d\nu \end{aligned} \tag{6.23}$$

The first term, which contains the the static dielectric permittivities of the three media $\varepsilon_1$, $\varepsilon_2$, and $\varepsilon_3$, represents the Keesom plus the Debye energy. It plays an important role for forces in water since water molecules have a strong dipole moment. Usually, however, the second term dominates in Eq. (6.23). The dielectric permittivity is not a constant but it depends on the frequency of the electric field. The static dielectric permittivities are the values of this dielectric function at zero frequency. $\varepsilon_1(i\nu)$, $\varepsilon_2(i\nu)$, and $\varepsilon_3(i\nu)$ are the dielectric permittivities at imaginary frequencies $i\nu$, and $\nu_1 = 2\pi k_B T/h = 3.9 \times 10^{13}$ Hz at 25°C. This corresponds to a wavelength of 760 nm, which is the optical regime of the spectrum. The energy is in the order of electronic states of the outer electrons.

In order to calculate the Hamaker constant the dielectric properties of all three materials need to be known. For frequencies starting in the visible range the dielectric permittivity can be described by

$$\varepsilon(i\nu) = 1 + \frac{n^2 - 1}{1 + \nu^2/\nu_e^2} \tag{6.24}$$

[5] A complete equation of the van der Waals force, which also contains retarded parts, is described in Ref. [119].

Here, $n$ is the refractive index and is the $\nu_e$ the mean ionization frequency of the material. Typically this is $\nu_e \approx 3{\times}10^{15}$ Hz. If the absorption frequencies of all three materials are assumed to be the same, we obtain the following approximation for the non-retarded Hamaker constant:

$$\begin{aligned} A_H &\approx \frac{3}{4}k_B T \cdot \left(\frac{\varepsilon_1 - \varepsilon_3}{\varepsilon_1 + \varepsilon_3}\right) \cdot \left(\frac{\varepsilon_2 - \varepsilon_3}{\varepsilon_2 + \varepsilon_3}\right) \\ &+ \frac{3h\nu_e}{8\sqrt{2}} \cdot \frac{\left(n_1^2 - n_3^2\right) \cdot \left(n_2^2 - n_3^2\right)}{\sqrt{n_1^2 + n_3^2} \cdot \sqrt{n_2^2 + n_3^2} \cdot \left(\sqrt{n_1^2 + n_3^2} + \sqrt{n_2^2 + n_3^2}\right)} \end{aligned} \tag{6.25}$$

Here, we used the fact that the major contribution to the Hamaker constant comes from frequencies in the visible or UV. The refractive indices of the materials are $n_1$, $n_2$, and $n_3$. Values of $\varepsilon$, $n$, and $\nu_e$ for different materials are listed in Table 6.2.

■ **Example 6.6.** Calculate the Hamaker constant for the interaction of amorphous silicon oxide ($SiO_2$) with silicon oxide across water at 20°C. According to Table 6.2 we insert $\varepsilon = 78.5$ and $n = 1.33$ for water, $\varepsilon = 3.82$ and $n = 1.46$ for silicon oxide, and $\nu_e = 3.4 \times 10^{15}$ Hz for the mean absorption frequency, into Eq. (6.25):

$$\begin{aligned} A_H &= 0.309 \times 10^{-20}\,\mathrm{J} \cdot \left(\frac{3.82 - 78.5}{3.82 + 78.5}\right)^2 \\ &+ 59.7 \times 10^{-20}\,\mathrm{J} \cdot \frac{\left(1.46^2 - 1.33^2\right)}{\left(1.46^2 + 1.33^2\right) \cdot \left(2\sqrt{1.46^2 + 1.33^2}\right)} \\ &= 0.280 \times 10^{-20}\,\mathrm{J} + 1.41 \times 10^{-20}\,\mathrm{J} = 1.69 \times 10^{-20}\,\mathrm{J} \end{aligned}$$

Using full spectroscopic results and Eq. (6.23), leads to somewhat lower Hamaker constants (Table 6.3).

Equation (6.25) not only allows us to calculate the Hamaker constant, it also allows us to easily predict whether we can expect attraction or repulsion. An attractive van der Waals force corresponds to a positive sign of the Hamaker constant, repulsion corresponds to a negative Hamaker constant. Van der Waals forces between similar materials are always attractive. This can easily be deduced from the last equation: for $\varepsilon_1 = \varepsilon_2$ and $n_1 = n_2$ the Hamaker constant is positive, which corresponds to an attractive force. If two different media interact across vacuum ($\varepsilon_3 = n_3 = 1$), or practically a gas, the van der Waals force is also attractive. Van der Waals forces between different materials across a condensed phase can be repulsive. Repulsive van der Waals forces occur, when medium 3 is more strongly attracted to medium 1 than medium 2. Repulsive forces were, for instance, measured for the interaction of silicon nitride with silicon oxide in diiodomethane [121]. Repulsive van der Waals forces can also occur across thin films on solid surfaces. In the case of thin liquid films on solid surfaces there is often a repulsive van der Waals force between the solid–liquid and the liquid–gas interface [122].

An important example of a repulsive van der Waals force is the force between a solid particle interacting in water with an air bubble. This is a typical situation in flotation, where air bubbles are used to extract mineral particles from an aqueous dispersion (see Section 7.6.1). For some materials the van der Waals force between the solid–liquid and the liquid–vapor

**Table 6.2:** Dielectric permittivity $\varepsilon$, refractive index $n$, and main absorption frequency in the UV $\nu_e$ for various solids, liquids, and polymers at 20°C (Refs. [109, 128, 129], handbooks, and own measurements.).

| Material | $\varepsilon$ | $n$ | $\nu_e$ ($10^{15}$ Hz) |
|---|---|---|---|
| $Al_2O_3$ (alumina) | 9.3–11.5 | 1.75 | 3.2 |
| C (diamond) | 5.7 | 2.40 | 2.7 |
| $CaCO_3$ (calcium carbonate, average) | 8.2 | 1.59 | 3.0 |
| $CaF_2$ (fluorite) | 6.7 | 1.43 | 3.8 |
| $KAl_2Si_3AlO_{10}(OH)_2$ (muscovite mica) | 5.4 | 1.58 | 3.1 |
| KCl (potassium chloride) | 4.4 | 1.48 | 2.5 |
| NaCl (sodium chloride) | 5.9 | 1.53 | 2.5 |
| $Si_3N_4$ (silicon nitride, amorphous) | 7.4 | 1.99 | 2.5 |
| $SiO_2$ (quartz) | 4.3–4.8 | 1.54 | 3.2 |
| $SiO_2$ (silica, amorphous) | 3.82 | 1.46 | 3.2 |
| $TiO_2$ (titania, average) | 114 | 2.464 | 1.2 |
| ZnO (zinc oxide) | 11.8 | 1.910 | 1.4 |
| Acetone | 20.7 | 1.359 | 2.9 |
| Chloroform | 4.81 | 1.446 | 3.0 |
| $n$-Hexane | 1.89 | 1.38 | 4.1 |
| $n$-Octane | 1.97 | 1.41 | 3.0 |
| $n$-Hexadecane | 2.05 | 1.43 | 2.9 |
| Ethanol | 25.3 | 1.361 | 3.0 |
| 1-Propanol | 20.8 | 1.385 | 3.1 |
| 1-Butanol | 17.8 | 1.399 | 3.1 |
| 1-Octanol | 10.3 | 1.430 | 3.1 |
| Toluene | 2.38 | 1.497 | 2.7 |
| 1Water | 78.5 | 1.333 | 3.6 |
| Polyethylene | 2.26–2.32 | 1.48–1.51 | 2.6 |
| Polystyrene | 2.49–2.61 | 1.59 | 2.3 |
| Poly(vinyl chloride) | 4.55 | 1.52–1.55 | 2.9 |
| Poly(tetrafluoroethylene) | 2.1 | 1.35 | 4.1 |
| Poly(methyl methacrylate) | 3.12 | 1.50 | 2.7 |
| Poly(ethylene oxide) | | 1.45 | 2.8 |
| Poly(dimethyl siloxane) | 2.6–2.8 | 1.4 | 2.8 |
| Nylon 6 | 3.8 | 1.53 | 2.7 |
| Bovine serum albumin | 4.0 | | 2.4-2.8 |

interfaces is repulsive and a stable water film is formed (p. 87 in Ref. [123]). This prevents flotation.

The above analysis applies to insulating materials. For electrically conductive materials such as metals the dielectric constant is infinite and Eq. (6.24) does not apply. In this case we

can approximate the dielectric permittivity of the metal by

$$\varepsilon(i\nu) = 1 + \frac{\nu_e^2}{\nu^2}, \tag{6.26}$$

where $\nu_e$ is the so called plasma frequency of the electron gas; it is typically $5 \times 10^{15}$ Hz. Inserting Eq. (6.26) into Eq. (6.23) leads to the approximate Hamaker constant

$$A_H \approx \frac{3}{16\sqrt{2}} h\nu_e \approx 4 \times 10^{-19}\,\mathrm{J} \tag{6.27}$$

for two metals interacting across a vacuum. Thus, the Hamaker constants of metals and metal oxides can be up to an order of magnitude higher than those of non-conducting media.

In Table 6.3 non-retarded Hamaker constants are listed for different material combinations. Hamaker constants, calculated from spectroscopic data, are found in many publications [124–128]. A review is given in Ref. [129].

**Table 6.3:** Hamaker constants for medium 1 interacting with medium 2 across medium 3. The variation in calculated Hamaker constants is due to the fact that different dielectric functions were used in Eq. (6.23). Results were partially taken from Refs. [124, 126–130].

| Medium 1 | Medium 3 | Medium 2 | $A_H$ calc. ($10^{-20}$ J) | $A_H$ exp. ($10^{-20}$ J) |
|---|---|---|---|---|
| Au/Ag/Cu | Vacuum | Au/Ag/Cu | 20–50 | |
| Mica | Vacuum | Mica | 7.0 | 10–13.5 |
| $Al_2O_3$ | Vacuum | $Al_2O_3$ | 14.5-15.2 | |
| $SiO_2$ | Vacuum | $SiO_2$ | 6.4–6.6 | 5–6 |
| $Si_3N_4$ | Vacuum | $Si_3N_4$ | 16.2–17.4 | |
| $TiO_2$ | Vacuum | $TiO_2$ | 14.3–17.3 | |
| Perfluorocarb. | Vacuum | Perfluorocarb. | 3.4–6.0 | |
| Carbonhydr. | Vacuum | Carbonhydr. | 2.6–3.0 | |
| Au/Ag/Cu | Water | Au/Ag/Cu | 10–13 | |
| Mica | Water | Mica | 0.29 | 2.2 |
| $Al_2O_3$ | Water | $Al_2O_3$ | 2.8–4.7 | 6.7 |
| $SiO_2$ | Water | $SiO_2$ | 0.16-1.51 | |
| $Si_3N_4$ | Water | $Si_3N_4$ | 4.6–5.9 | 2–8 |
| $TiO_2$ | Water | $TiO_2$ | 5.4–6.0 | 4–8 |
| Perfluorocarb. | Water | Perfluorocarb. | 0.36-0.74 | |
| Carbonhydr. | Water | Carbonhydr. | 0.39–0.44 | 0.3–0.6 |
| Polystyrene | Water | Polystyrene | 0.9–1.3 | |
| $SiO_2$ | Water | Air | -1.0 | |
| BSA (Albumin) | Water | $SiO_2$ | 0.7 | |

In many applications we are interested in the van der Waals force across an aqueous medium. Then, an important question concerns the effect of dissolved ions. Ions hinder

the water molecules in their hydration shell from orienting in an external electric field. The first term in the equation of the Hamaker constant is affected. In addition, the salt concentration is often much higher on surfaces than in the bulk phase. As a consequence, the dielectric constant can be smaller than in the bulk phase.

■ **Example 6.7.** For the interaction of lipid bilayers across a layer of water, a Hamaker constant of $7.5\times10^{-21}$ J is calculated. A value of only $3\times10^{-21}$ J was measured. One reason is probably a reduction of the first term by the presence of ions [131].

From Eq. (6.23) a useful approximation can be derived:

$$A_{132} \approx \sqrt{A_{131} \cdot A_{232}} \tag{6.28}$$

If we know the Hamaker constant of material 1 interacting across medium 3 with itself, $A_{131}$, and we know the Hamaker constant of material 2 interacting across medium 3 with itself, $A_{232}$, then we can estimate the Hamaker constant for the interaction between material 1 with material 2 across medium 3, $A_{132}$.

### 6.2.3 Surface energy and Hamaker constant

In order to calculate the surface energy of molecular crystals, we imagine the following experiment: A crystal is cleaved in two parts which are separated by an infinite distance (Fig. 6.3). The work required per unit area is $w = A_H/12\pi D_0^2$, where $D_0$ is the distance between two atoms. Upon cleavage, two fresh surfaces are formed. With the surface energy $\gamma_S$ the work required is $2\gamma_S$. Equating the results leads to

$$\gamma_S = \frac{A_H}{24\pi D_0^2} \tag{6.29}$$

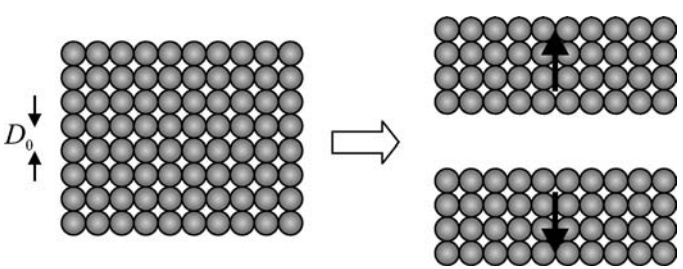



**Figure 6.3:** Cleaving a molecular crystal to calculate the surface energy of a solid.

■ **Example 6.8.** Helium [109]. As an interatomic distance often a value of 1.6 Å is used. The Hamaker constant of helium-vacuum-helium is $5.7 \times 10^{-22}$ J. Calculating the surface energy with Eq. (6.29) leads to 0.29 mJ/m$^2$. This value is in good agreement with measured values for liquid helium of 0.12–0.35 mJ/m$^2$.

For Teflon we calculate a surface energy of 16–28 mJ/m$^2$ [127] using an atomic spacing of 1.7 Å and a Hamaker constant of $3.4$–$6.0\times10^{-20}$ J.

## 6.3 Concepts for the description of surface forces

### 6.3.1 The Derjaguin approximation

In the previous section we calculated the van der Waals force between two spheres and between two planar surfaces. What if the two interacting bodies do not have such a simple geometry? We could try to do an integration similar to what was described for the two spheres. This, however, might be very difficult and lead to long expressions. The Derjaguin[6] approximation is a simple way to overcome this problem. The Derjaguin approximation relates the energy per unit area between two planar surfaces $w$ which are separated by a gap of width $x$ to the energy between two bodies of arbitrary shape $W$ which are at a distance $D$:

$$W(D) = \int_D^\infty w(x) dA \tag{6.30}$$

The integration runs over the entire surface of the solid. Please note that here $A$ is the cross-sectional area. Often we have to deal with rotational-symmetric configurations. Then it is reasonable to integrate in cylindrical coordinates:

$$W(D) = 2\pi \int_0^\infty w\left(x(r)\right) \cdot r \cdot dr \tag{6.31}$$

In many cases the following expression is more useful:

$$W(D) = \int_D^\infty w\left(x\right) \frac{dA}{dx} dx \tag{6.32}$$

The approximation is only valid if the characteristic decay length of the surface force is small in comparison to the curvature of the surfaces. Approximation (6.30) is sometimes called the Derjaguin approximation in honor of Derjaguin's work. He used this approach to calculated the interaction between two ellipsoids [132].

■ **Example 6.9.** Calculate the van der Waals force between a cone with an opening angle $\alpha$ and a planar surface (Fig. 6.4). The cross-sectional area is given by

$$A = \pi \cdot \left[(x - D)\tan\alpha\right]^2 \qquad \text{for} \qquad x \geq D$$

which leads to

$$\frac{dA}{dx} = 2\pi \cdot \tan^2\alpha \cdot (x - D)$$

[6] Boris Vladimirovich Derjaguin, 1902–1994. Russian physicochemist, professor in Moscow.

For the force we can use a similar equation as for the energy:

$$\begin{aligned} F(D) &= \int_D^\infty f(x)\frac{dA}{dx}dx = -\int_D^\infty \frac{A_H}{6\pi x^3}\cdot 2\pi\cdot\tan^2\alpha\cdot(x-D)\cdot dx \\ &= -\frac{A_H\tan^2\alpha}{3}\cdot\int_D^\infty \frac{x-D}{x^3}dx \end{aligned}$$

$$\begin{aligned} F(D) &= -\frac{A_H\tan^2\alpha}{3}\cdot\left[-\frac{1}{x}+\frac{D}{2x^2}\right]_D^\infty = \frac{A_H\tan^2\alpha}{3}\cdot\left(-\frac{1}{D}+\frac{1}{2D}\right) \\ &= -\frac{A_H\tan^2\alpha}{6D} \end{aligned}$$

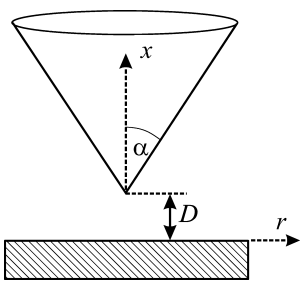



**Figure 6.4:** Interaction between a cone and a planar surface.

A special, but nevertheless important, case is the interaction between two identical spheres. It is important to understand the stability of dispersions. For the case of two spheres of equal radius $R$, the parameters $x$ and $r$ are related by (Fig. 6.5)

$$x(r) = D+2R-2\sqrt{R^2-r^2} \Rightarrow dx = \frac{2r}{\sqrt{R^2-r^2}}dr \Rightarrow 2rdr = \sqrt{R^2-r^2}\cdot dx \quad (6.33)$$

If the range of the interaction is substantially smaller than $R$, then we only need to consider the outer caps of the two spheres and only the contributions with a small $r$ are effective. We can simplify

$$2rdr = \sqrt{R^2-r^2}\cdot dx \approx Rdx \quad (6.34)$$

and integral (6.32) becomes

$$W(D) = \pi R\cdot\int_D^\infty w(x)\,dx \quad (6.35)$$

From the potential energy we can calculate the force between two spheres:

$$F = -\frac{dW}{dD} = -\pi R\cdot\frac{d}{dD}\left(\int_D^\infty w(x)\,dx\right) = \pi R\cdot w(D) \quad (6.36)$$

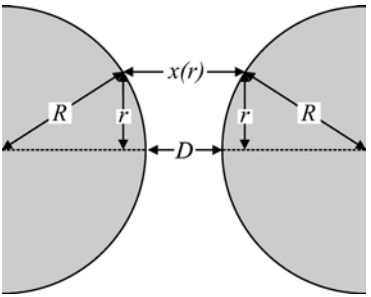



**Figure 6.5:** Calculating the interaction between two spheres with Derjaguin's approximation.

since $w(\infty) = 0$. Thus, assuming that the range of the interaction is substantially smaller than $R$, there is a simple relationship between the potential energy per unit area and the force between two spheres.

As one example we calculate the van der Waals force between two identical spheres from the van der Waals energy per unit area $w = -A_H/12\pi x^2$ (Eq. 6.17). Using Derjaguin's approximation we can directly write

$$F = -\frac{A_H R}{12D^2} \tag{6.37}$$

which agrees with Eq. (6.21) for $R_1 = R_2 = R$.

Equation (6.36) refers to the interaction between two spheres. For a sphere which is at a distance $D$ from a planar surface we get a similar relation:

$$F = 2\pi R \cdot w(D) \tag{6.38}$$

White [133] extended the approximation to solids with arbitrary shape.

Derjaguin's approximation has a fundamental consequence. In general, the force or energy between two bodies depends on the shape, on the material properties, and on the distance. Now it is possible to divide the force (or energy) between two solids into a purely geometrical factor and into a material and distance-dependent term $w(x)$. Thus it is possible to describe the interaction independent of the geometry. This also gives us a common reference, which is $w(x)$. In the following we only discuss $w(x)$.

### 6.3.2 The disjoining pressure

The term "disjoining pressure" was introduced in 1936 by Derjaguin [134]. The disjoining pressure $\Pi$ is equal to the difference between the pressure within a film between two surfaces and the pressure in the bulk phase (Fig. 6.6). It is defined as the change in Gibbs free energy with distance and per unit area at constant cross-sectional area, temperature, and volume:

$$\Pi = -\frac{1}{A} \cdot \left.\frac{\partial G}{\partial x}\right|_{A,T,V} \tag{6.39}$$

The film is in equilibrium with the bulk phase.

The concept of disjoining pressure is not in contradiction to the formalism of surface forces. It is sometimes more useful to think in terms of disjoining pressure. For example, if

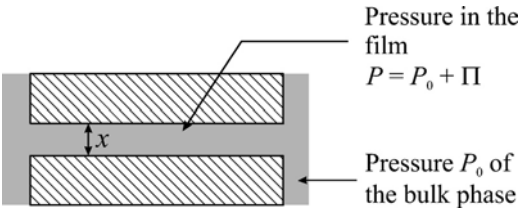



**Figure 6.6:** The disjoining pressure between two parallel plates.

the stability of a liquid foam film is examined, it is more appropriate to think in terms of an increased pressure in the film rather than of the fact that this pressure is caused by the force between two liquid–air interfaces.

## 6.4 Measurement of surface forces

The development of the theory of van der Waals forces stimulated an interest in measuring forces between surfaces to verify the theory (review: Ref. [135]). Early direct force measurements were made between two polished glass bodies [136]. One glass body was fixed, the other was mounted to a spring. The distance between the two glass surfaces and the deflection of the spring were measured. Multiplication of the deflection by the spring constant, gave the force. Using these simple early devices and, after overcoming severe problems, it was possible to verify the theoretically predicted van der Waals force for glass interacting with glass across a gaseous medium [137, 138]. Derjaguin, Rabinovich, and Churaev measured the force between two thin metal wires. In this way they could determine the van der Waals force for metals [139].

In these early measurements several problems became obvious. The minimal distance accessible ($\approx$ 20 nm) was rather limited and experiments could not be done in a liquid environment. One inherent problem of all techniques used to measure surface forces is surface roughness. The surface roughness over the interacting areas limits the distance resolution and the accuracy of how well zero distance (contact) is defined. Practically, zero distance is the distance of closest approach because if a protrusion is sticking out of one surface, the other surface cannot approach any closer. Two methods can be used to reduce this problem of surface roughness. Either atomically smooth surfaces are chosen, or the interacting areas are reduced. The first approach was realized with the development of the surface forces apparatus while the advantage of small interacting areas could be utilized with the invention of the atomic force microscope.

The development of the **surface forces apparatus** (SFA) [140, 141] was a big step forward because it allowed one to measure directly the force law in liquids and vapors at Ångström resolution level [142]. The SFA contains two crossed atomically smooth mica cylinders of roughly 1 cm radius between which the interaction forces are measured (Fig. 6.7). One mica cylinder is mounted to a piezoelectric translator, with which this translator the distance is adjusted. The other mica surface is mounted to a spring of known and adjustable spring constant. The separation between the two surfaces is measured by use of an optical technique using multiple beam interference fringes. Knowing the position of one cylinder and the separation from the surface of the second cylinder, the deflection of the spring and the force can be calculated.

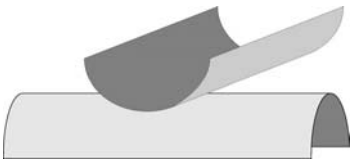


**Figure 6.7:** Schematic of two crossed mica cylinders used in a surface forces apparatus.

The second device with which surface forces can be measured directly and relatively universally is the **atomic force microscope** (AFM) sometimes also called the scanning force microscope (Fig. 6.8) [143, 144]. In the atomic force microscope we measure the force between a sample surface and a microfabricated tip, placed at the end of an about 100 $\mu$m long and 0.4–10 $\mu$m thick cantilever. Alternatively, colloidal particles are fixed on the cantilever. This technique is called the "colloidal probe technique". With the atomic force microscope the forces between surfaces and colloidal particles can be directly measured in a liquid [145, 146]. The practical advantage is that measurements are quick and simple. Even better, the interacting surfaces are substantially smaller than in the surface forces apparatus. Thus the problem of surface roughness, deformation, and contamination, is reduced. This again allows us to examine surfaces of different materials.

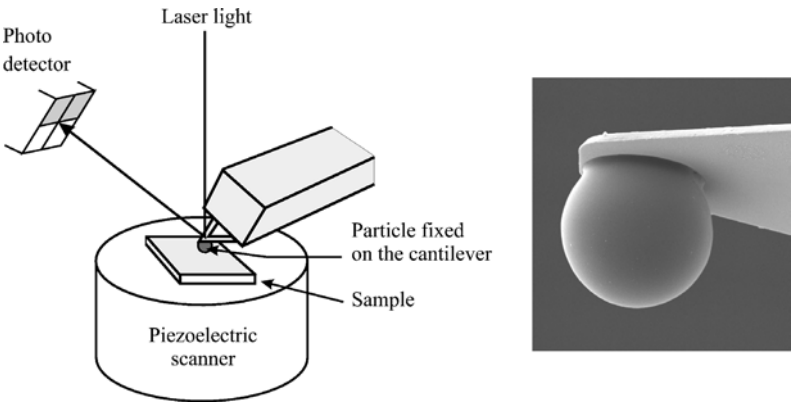



**Figure 6.8:** Schematic of an atomic force microscope (left) and an example of a colloidal probe. The glass sphere is sintered to the end of an AFM cantilever has a diameter of $\approx 10\,\mu$m.

Another important, although less direct, technique for measuring forces between macromolecules or lipid bilayers is the **osmotic stress method** [147, 148]. A dispersion of vesicles or macromolecules is equilibrated with a reservoir solution containing water and other small solutes, which can freely exchange with the dispersion phase. The reservoir also contains a polymer which cannot diffuse into the dispersion. The polymer concentration determines the osmotic stress acting on the dispersion. The spacing between the macromolecules or vesicles is measured by X-ray diffraction. In this way one obtains pressure-versus-distance curves. The osmotic stress method is used to measure interactions between lipid bilayers, DNA, polysaccharides, proteins, and other macromolecules [149]. It was particularly successful in studying the hydration force (see below) between lipid bilayers and biological macromolecules.

All techniques mentioned so far are mainly used to study the force between solid surfaces. In many applications one is interested in the disjoining pressure between liquid–liquid or liquid–gas interfaces, such as those found in foams and emulsions. One such technique is described in Section 12.5.3.

## 6.5 The electrostatic double-layer force

In Chapter 5 we learned that, in water, most surfaces bear an electric charge. If two such surfaces approach each other and the electric double layers overlap, an electrostatic double-layer force arises. This electrostatic double-layer force is important in many natural phenomena and technical applications. It for example stabilizes dispersions.[7]

The interaction between two charged surfaces in liquid depends on the surface charge. Here, we only consider the linear case and assume that the surface potentials are low. If we had to use the nonlinear Poisson–Boltzmann theory, the calculations would become substantially more complex. In addition, only monovalent salts are considered. An extension to other salts can easily be made.

### 6.5.1 General equations

When calculating the Gibbs free energy per unit area $w(x)$ two approaches can be used. Either the change in Gibbs free energy of the two double layers during the approach is calculated, or the disjoining pressure in the gap is determined. Both approaches lead to the same result. Depending upon conditions and on personal preference, one or the other method is more suitable. Hogg, Healy and Fuerstenau used the first condition in order to calculate the energy between two spheres with constant, but different, surface potentials [152]. Parsegian and Gingell chose to determine the osmotic pressure for two different surfaces with different boundary conditions [153].

We start with the first approach and calculate the change in Gibbs free energy for two approaching double layers. The Gibbs free energy of one isolated electric double layer per unit area is (Eq. 4.35)

$$-\int_0^{\psi_0} \sigma' d\psi_0' \tag{6.40}$$

For two homogeneous double layers, which are infinitely separated, the Gibbs free energy per unit area is twice this value:

$$g^\infty = -2\int_0^{\psi_0} \sigma' d\psi_0' \tag{6.41}$$

[7] Please note that the electrostatic double-layer force is fundamentally different from the Coulomb force. For example, if we consider two identical spherical particles of radius $R$ you cannot take Eq. (6.1), insert the total surface charge as $Q_1$ and $Q_2$, use the dielectric permittivity of water and expect to get a reasonable result. The main differences are the free charges (ions) in solution. They screen the electrostatic field emanating from the surfaces.

If the two surfaces approach each other up to a distance $x$, the Gibbs free energy changes. Now, surface charge and potential depend on the distance:

$$g(x) = -2 \int_0^{\psi_0(x)} \sigma'(x) d\psi_0' \tag{6.42}$$

The Gibbs free interaction energy per unit area is

$$w(x) = \Delta g = g(x) - g^\infty \tag{6.43}$$

For the force per unit area we get

$$\Pi = -\frac{d\Delta g}{dx} \tag{6.44}$$

Scientist like to use equations and the aim of a theory is often to derive an equation. Whole theories can be built around a single equation. Equations are important because they allow precise, quantitative predictions. Dependencies become obvious. Practically, they can also lead us to a better understanding of a process since they guide us to its cause. As one example we ask, why do charged surfaces interact? In order to get an intuitive understanding we consider what happens to surface charges when two surfaces are brought closer, holding the potential constant. Remember: surface charge and surface potential are related by (Eq. 4.26)

$$\sigma = -\varepsilon\varepsilon_0 \left.\frac{d\psi}{d\xi}\right|_{\xi=0} \tag{6.45}$$

The surface charge is (except for two constants) equal to the potential gradient $|d\psi/d\xi|$ at the surface. If two surfaces approach each other, the potential gradient decreases (Fig. 6.9). The surface charge density decreases accordingly. If, for instance, AgI particles are brought together, then $I^-$ ions are removed from the surface and the surface charge density decreases. During the approach of $SiO_2$ particles, previously dissociated $H^+$ ions bind again. Neutral hydroxide groups are formed and the negative surface charge becomes weaker. This reduction of the surface charge increases their Gibbs free energy (otherwise the double layer would not have formed). As a consequence the surfaces repel each other.

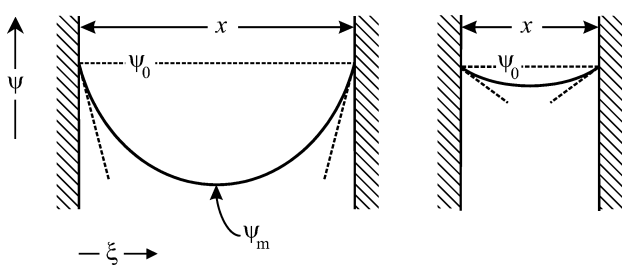



**Figure 6.9:** Change in the potential distribution when two parallel planar surfaces approach each other. The gap is filled with electrolyte solution.

At this point it is probably instructive to discuss the use of the symbols $D$, $x$, and $\xi$. $D$ is the shortest distance between two solids of arbitrary geometry. Usually we use $x$ for the thickness of the gap between two infinitely extended solids. For example, it appears in the Derjaguin approximation because there we integrate over many such hypothetical gaps. $\xi$ is a coordinate describing a position within the gap. At a given gap thickness $x$, the potential changes with $\xi$ (Fig. 6.9). $D$ is the distance between finite, macroscopic bodies.

Alternatively, the interaction can be calculated with the disjoining pressure:

$$w(x) = -\int_{\infty}^{x} \Pi(x')dx' \tag{6.46}$$

Using Poisson–Boltzmann theory we can derive a simple expression for the disjoining pressure. For the linear case (low potentials) and for a monovalent salt, the one-dimensional Poisson–Boltzmann equation (Eq. 4.9) is

$$ec_0\left(e^{\frac{e\psi}{k_BT}} - e^{-\frac{e\psi}{k_BT}}\right) - \varepsilon\varepsilon_0 \cdot \frac{d^2\psi}{dx^2} = 0 \tag{6.47}$$

The first term is sometimes written as $2ec_0 \cdot \sinh\left(e\psi/k_BT\right)$. Integration leads to

$$c_0k_BT\left(e^{\frac{e\psi}{k_BT}} + e^{-\frac{e\psi}{k_BT}}\right) - \frac{\varepsilon\varepsilon_0}{2} \cdot \left(\frac{d\psi}{dx}\right)^2 = P \tag{6.48}$$

So far $P$ is only an integration constant. As we see later it has a physical meaning: $P$ corresponds to the pressure in the gap. The first term describes the osmotic pressure caused by the increased number of particles (ions) in the gap. The second term, sometimes called the Maxwell stress term, corresponds to the electrostatic force caused by the electric field of one surface which affects charges on the other surface and vice versa.

To obtain the disjoining pressure we have to realise that the solution in the infinitely extended gap is in contact with an infinitely large reservoir (Fig. 6.6) As the force per unit area is $\Pi$ only the difference of the pressure inside the gap and the pressure in the reservoir is effective. Therefore the osmotic pressure in the reservoir $2k_BTc_0$ must be subtracted from $P$ in order to get the disjoining pressure: $\Pi = P - 2k_BTc_0$. Finally, for the force per unit area we obtain

$$\Pi = c_0k_BT \cdot \left(e^{\frac{e\psi}{k_BT}} + e^{-\frac{e\psi}{k_BT}} - 2\right) - \frac{\varepsilon\varepsilon_0}{2} \cdot \left(\frac{d\psi}{d\xi}\right)^2 \tag{6.49}$$

In order to determine the force in a specific situation, the potential must first be calculated. This is done by solving the Poisson–Boltzmann equation. In a second step, the force per unit area is calculated. It does not matter for which point $\xi$ we calculate $\Pi$, the value must be the same for every $\xi$.

In order to calculate the potential distribution in the gap we need not only the Poisson–Boltzmann equation, but in addition, boundary conditions must be specified. Two common types of boundary conditions are:

**Constant potential**: Upon approach of the two surfaces the surface potentials remain constant: $\psi(\xi = 0) = \psi_1$ and $\psi(\xi = x) = \psi_2$.

**Constant charge**: During approach the surface charge densities $\sigma_1$ and $\sigma_2$ are constant.

For large distances ($x \gg \lambda_D$) both boundary conditions lead to identical forces. At small distances the constant charge condition leads to more repulsive forces than the constant potential condition (Fig. 6.10).

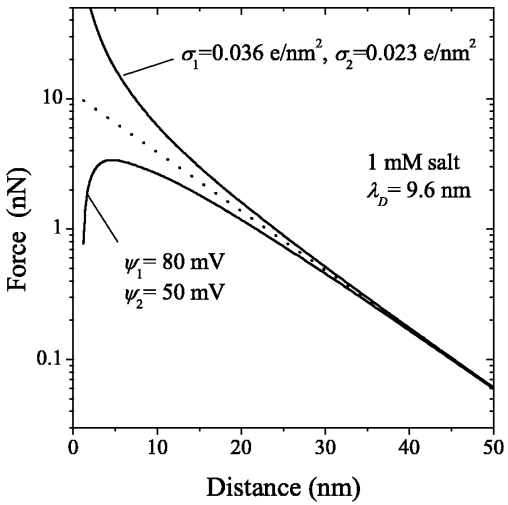



**Figure 6.10:** Electrostatic double-layer force between a sphere of $R = 3$ $\mu$m radius and a flat surface in water containing 1 mM monovalent salt. The force was calculated using the nonlinear Poisson–Boltzmann equation and the Derjaguin approximation for constant potentials ($\psi_1 = 80$ mV, $\psi_2 = 50$ mV) and for constant surface charge ($\sigma_1 = 0.0058\,\mathrm{Cm}^{-2} = 0.036\,\mathrm{enm}^{-2}$, $\sigma_2 = 0.0036\,\mathrm{Cm}^{-2} = 0.023\,\mathrm{enm}^{-2}$). The surface charge was adjusted by $\sigma_{1/2} = \varepsilon\varepsilon_0\psi_{1/2}/\lambda_D$ so that at large distances both lead to the same potential.

### 6.5.2 Electrostatic interaction between two identical surfaces

An important case is the interaction between two identical parallel surfaces of two infinitely extended solids. It is, for instance, important to understand the coagulation of sols. We can use the resulting symmetry of the electric potential to simplify the calculation. For identical solids the surface potential $\psi_0$ on both surfaces is equal. In between, the potential decreases (Fig. 6.9). In the middle the gradient must be zero because of the symmetry, i.e. $d\psi\left(\xi = x/2\right)/d\xi = 0$. Therefore, the disjoining pressure in the center is given only by the osmotic pressure. Towards the two surfaces, the osmotic pressure increases. This increase is, however, compensated by a decrease in the Maxwell stress term. Since in equilibrium the pressure must be the same everywhere, we have:

$$\Pi(x) = c_0 k_B T \left(e^{\frac{e\psi_m}{k_B T}} + e^{-\frac{e\psi_m}{k_B T}} - 2\right) \tag{6.50}$$

where $\psi_m$ is the electric potential in the middle. $\Pi$ depends on the gap thickness $x$, since $\psi_m$ changes with $x$. We can thus determine the force from the potential in the middle.

For low potentials we can further simplify this expression. Therefore we write the exponential functions in a series and neglect all terms higher than the quadratic one:

$$\begin{aligned}\Pi &= k_BTc_0\left(1+\frac{e\psi_m}{k_BT}+\left(\frac{e\psi_m}{k_BT}\right)^2+...+1-\frac{e\psi_m}{k_BT}+\left(\frac{e\psi_m}{k_BT}\right)^2\pm...-2\right)\\ &\approx \frac{2c_0e^2}{k_BT}\psi_m^2=\frac{\varepsilon\varepsilon_0}{\lambda_D^2}\cdot\psi_m^2 \qquad (6.51)\end{aligned}$$

It remains to find $\psi_m$. If the electric double layers of the two opposing surfaces overlap only slightly ($x \gg \lambda_D$), then we can approximate

$$\psi_m = 2\psi'(x/2) \qquad (6.52)$$

where $\psi'$ is the potential of an isolated double layer. For $\psi'$ we can use various exact functions. For low potentials we can use $\psi(\xi) = \psi_0 \cdot \exp(-\xi/\lambda_D)$, which leads to a repulsive force per area of

$$\Pi(x) = \frac{2\varepsilon\varepsilon_0}{\lambda_D^2}\cdot\psi_0^2\cdot e^{-x/\lambda_D} \qquad (6.53)$$

In order to calculate the Gibbs free interaction energy per unit area we still have to integrate:

$$w(x) = -\int_{\infty}^{x}\Pi(x')dx' = -\frac{2\varepsilon\varepsilon_0\psi_0^2}{\lambda_D^2}\int_{\infty}^{x}e^{-x'/\lambda_D dx'} = \frac{2\varepsilon\varepsilon_0\psi_0^2}{\lambda_D}\cdot e^{-x/\lambda_D} \qquad (6.54)$$

If we use expression (4.22) for $\psi'$, which is also valid at higher potentials, we get

$$w(x) = 64c_0k_BT\lambda_D\cdot\left(\frac{e^{e\psi_0/2k_BT}-1}{e^{e\psi_0/2k_BT}+1}\right)^2\cdot e^{-\frac{x}{\lambda_D}} \qquad (6.55)$$

This is sometimes written as

$$w(x) = 64c_0k_BT\lambda_D\cdot\tanh^2\left(\frac{e\psi_0}{4k_BT}\right)\cdot e^{-\frac{x}{\lambda_D}} \qquad (6.56)$$

Both expressions are identical. This is easily seen when remembering the definition of the tanh function:

$$\tanh z = \frac{e^z-e^{-z}}{e^z+e^{-z}}$$

Multiplication of both the numerator and denominator by $e^z$ leads to $\left(e^{2z}-1\right)/\left(e^{2z}+1\right)$.

### 6.5.3 The DLVO theory

It has been known for more than 100 years that many aqueous dispersions precipitate upon addition of salt. Schulze and Hardy observed that most dispersions precipitate at concentrations of 25–150 mM of monovalent counterions [154, 155]. For divalent ions they found far

smaller precipitation concentrations of 0.5–2 mM. Trivalent counterions lead to precipitation at even lower concentrations of 0.01–0.1 mM. For example, gold colloids are stable in NaCl solution, as long as the NaCl concentration does not exceed 24 mM. If the solution contains more NaCl, then the gold particles aggregate and precipitate. The appropriate concentrations for $KNO_3$, $CaCl_2$, and $BaCl_2$ are 23 mM, 0.41 mM and 0.35 mM [156].

This coagulation can be understood as follows. The gold particles are negatively charged and repel each other. With increasing salt concentration the electrostatic repulsion decreases. The particles, which move around thermally, frequently have a higher chance of approaching each other to a few Ångströms. Then the van der Waals attraction causes them to aggregate. Since divalent counterions weaken the electrostatic repulsion more effectively than monovalent counterions, only small concentrations of $CaCl_2$ and $BaCl_2$ are necessary for coagulation.

Roughly 60 years ago Derjaguin, Landau, Verwey, and Overbeek developed a theory to explain the aggregation of aqueous dispersions quantitatively [66, 157, 158]. This theory is called DLVO theory. In DLVO theory, coagulation of dispersed particles is explained by the interplay between two forces: the attractive van der Waals force and the repulsive electrostatic double-layer force. These forces are sometimes referred to as DLVO forces. Van der Waals forces promote coagulation while the double layer-force stabilizes dispersions. Taking into account both components we can approximate the energy per unit area between two infinitely extended solids which are separated by a gap $x$:

$$w(x) = 64c_0k_BT\lambda_D \cdot \left(\frac{e^{e\psi_0/2k_BT} - 1}{e^{e\psi_0}/2k_BT + 1}\right)^2 \cdot e^{-x/\lambda_D} - \frac{A_H}{12\pi x^2} \tag{6.57}$$

This is sometimes written as

$$w(x) = 64c_0k_BT\lambda_D \cdot \tanh^2\left(\frac{e\psi_0}{4k_BT}\right) \cdot e^{-x/\lambda_D} - \frac{A_H}{12\pi x^2} \tag{6.58}$$

Figure 6.11 shows the interaction energy between two identical spherical particles calculated with the DLVO theory. In general, it can be described by a very weak attraction at large distances (secondary energy minimum), an electrostatic repulsion at intermediate distances, and a strong attraction at short distances (primary energy minimum). At different salt concentration the three regimes are more or less pronounced and sometimes even completely missing. At low and intermediate salt concentration the repulsive electrostatic barrier prevents the particles from aggregating. With increasing salt concentration the repulsive energy barrier decreases. At low salt concentration the energy barrier is so high that particles in a dispersion have practically no chance of gaining enough thermal energy to overcome it. At high salt concentration this energy barrier is drastically reduced and the van der Waals attraction dominates. This leads to precipitation. In addition, the surface potential usually decreases with increasing salt concentration thus lowering the energy barrier even more (this effect was not taken into account in Fig. 6.11).

For small distances DLVO theory predicts that the van der Waals attraction always dominates. Please remember: the van der Waals force between identical media is always attractive irrespective of the medium in the gap. Thus thermodynamically, or after long periods of time, we expect all dispersions to precipitate. Once in contact, particles should not separate again, unless they are strongly hit by a third object and gain a lot of energy.

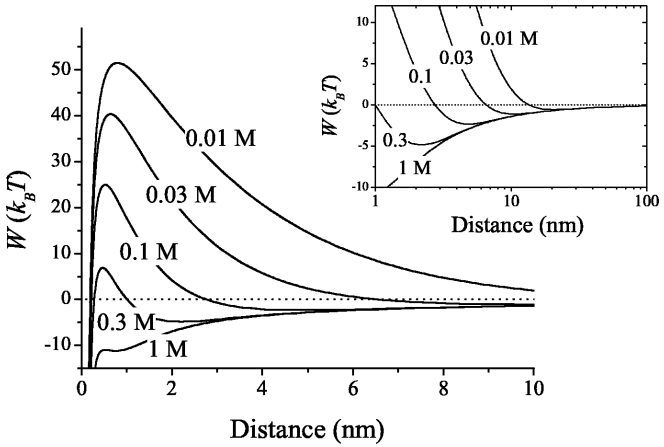



**Figure 6.11:** Gibbs free interaction energy (in units of $k_BT$) versus distance for two identical spherical particles of $R = 100$ nm radius in water, containing different concentrations of monovalent salt. The calculation is based on DLVO theory using Eqs. (6.57) and (6.32). The Hamaker constant was $A_H = 7 \times 10^{-21}$ J, the surface potential was set to $\psi_0 = 30$ mV. The insert shows the weak attractive interaction (secondary energy minimum) at very large distances.

A closer look at the interaction at large distances shows the weak attractive energy. This secondary energy minimum can lead to a weak, reversible coagulation without leading to direct molecular contact between the particles.

For many systems this is indeed observed. There are, however, important exceptions. One such exception is the swelling of clay [159–161]. In the presence of water or even water vapor, clay swells even at high salt concentrations. This cannot be understood based on DLVO theory . To understand phenomena liken the swelling of clay we have to consider the molecular nature of the solvent molecules involved.

## 6.6 Beyond DLVO theory

For large separations, the force between two solid surfaces in a fluid medium can usually be described by continuum theories such as the van der Waals and the electrostatic double-layer theory. The individual nature of the molecules involved, their discrete size, shape, and chemical nature was neglected. At surface separations approaching molecular dimensions continuum theory breaks down and the discrete molecular nature of the liquid molecules has to be taken into account.

### 6.6.1 The solvation force and confined liquids

Often the liquid structure close to an interface is different from that in the bulk. For many fluids the density profile normal to a solid surface oscillates about the bulk density with a periodicity of about one molecular diameter, close to the surface. This region typically extends

over a few molecular diameters. In this range the molecules are ordered in layers. Hints about this structure came from simulations and theory [162–165] while direct experimental proof is relatively recent [166, 167].

When two such surfaces approach each other, layer after layer is squeezed out of the closing gap (Fig. 6.12). Density fluctuations and the specific interactions then cause an exponentially decaying periodic force; the periodic length corresponds to the thickness of each layer. Such forces were termed **solvation forces** because they are a consequence of the adsorption of solvent molecules to solid surfaces [168]. Periodic solvation forces across confined liquids were first predicted by computer simulations and theory [168–171]. In this case, however, the experimental proof came only few years afterwards using the surface forces apparatus [172, 173]. Solvation forces are not only an important factor in the stability of dispersions. They are also important for analyzing the structure of confined liquids.

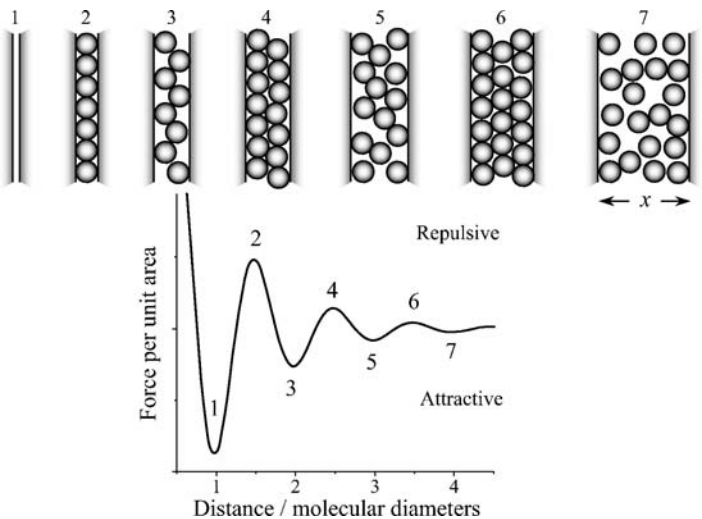



**Figure 6.12:** Schematic figure of the structure of a simple liquid confined between two parallel walls. The order changes drastically depending on distance, which results in a periodic force.

Solvation forces are often well described by an exponentially decaying cosine function of the form

$$f(x) = f_0 \cdot \cos\left(\frac{2\pi x}{d_0}\right) \cdot e^{-\frac{x}{x_0}} \tag{6.59}$$

Here, $f$ is the force per unit area, $f_0$ is the force extrapolated to $x = 0$, $d_0$ is the layer thickness, which in the case of simple liquids is equal to the molecular diameter, and $x_0$ is the characteristic decay length.

■ **Example 6.10.** The force across confined propanol is shown in Fig. 6.13. Clearly, at molecular separations the solvation force arising from the ordering of alcohol molecules dominates. The van der Waals force is plotted as a continuous line.

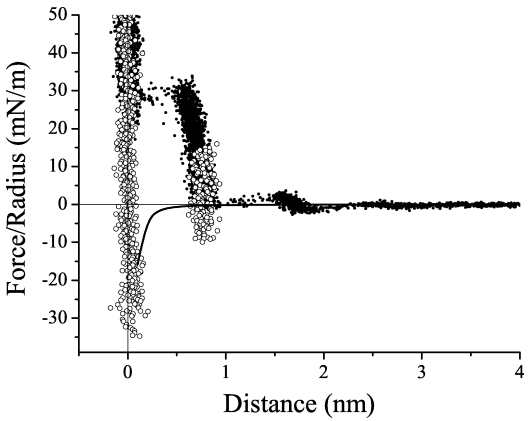



**Figure 6.13:** Normalized force between a microfabricated silicon nitride tip of an atomic force microscope and a planar mica surface in 1-propanol at room temperature [174]. The tip had a radius of curvature of $R \approx 50$ nm. The different symbols were recorded during approach (filled circles) and retraction (open circles) of the tip.

### 6.6.2 Non DLVO forces in an aqueous medium

Non DLVO forces in water deserve a special subchapter because they are important and far from being understood. They are important because water is the universal solvent in nature. Also, in more and more industrial processes water is used instead of organic solvent since it is harmless to the environment.

When two hydrophilic surfaces are brought into contact, repulsive forces of about 1 nm range have been measured in aqueous electrolyte between a variety of surfaces: Clays, mica, silica, alumina, lipids, DNA, and surfactants. Because of the correlation with the low (or negative) energy of wetting of these solids with water, the repulsive force has been attributed to the energy required to remove the water of hydration from the surface, or the surface adsorbed species, presumably because of strong charge–dipole, dipole–dipole or H-bonding interactions. These forces were termed **hydration forces** (reviews: [149, 175–177]).

Up to now the origin of hydration forces is not clear and several effects are discussed. Certainly the fact that one layer of water molecules is bound to the solid surfaces is important. The hydration force, however, extends over more than only two water layers. Israelachvili and Wennerstöm point out that the effect of the first water layer should not even be called a hydration force because it is caused by the interaction between water molecules and the solid surface and not by water–water interactions [175]. In a classical paper Marcelja and Radic proposed an elegant theory to explain the short-range repulsion by a modification of water structure near hydrophilic surfaces [178]. Modern theories take additional effects into account. In fact, short-range monotonically repulsive forces observed between inorganic surfaces are probably not only due to structured water layers propagated away from the surfaces, but to the osmotic effect of hydrated ions which are electrostatically trapped between two approaching surfaces [179]. This is supported by the observation that the hydration force is

strongly influenced by the ion concentration. In particular, cations, which are well known to have a shell of water molecules around them, tend to increase the hydration repulsion.

It is quite possible that several effects contribute to the short-range repulsion. This is especially likely for the interaction between flexible surfaces such as lipid bilayers [180, 181]. Molecular-scale fluctuations of hydrocarbon chains and a steric repulsion between mobile head groups may significantly contribute [182]. In lipid bilayers the individual lipid molecules are not fixed to a certain position but they thermally jump up and down by several Ångströms. If two such bilayers approach each other, this fluctuation is hindered. As a result the entropy of the lipid molecules decreases, their Gibbs free energy increases, and the two bilayers repel each other. Molecular scale corrugations can cause a short-range repulsion [183].

Between hydrophobic surfaces a completely different interaction is observed. Hydrophobic surfaces attract each other [184]. This attraction is called **hydrophobic interaction**. The first direct evidence that the interaction between solid hydrophobic surfaces is stronger than the van der Waals attraction was provided by Pashley and Israelachvili [185, 186]. With the surface force apparatus they observed an exponentially decaying attractive force between two mica surfaces with an adsorbed monolayer of the cationic surfactant cetyltrimethylammonium bromide (CTAB). Since then the hydrophobic force has been investigated by different groups and its existence is now generally accepted [189]. The origin of the hydrophobic force is, however, still under debate.

Usually two components of the attraction are observed [187]. One is short-range and decays roughly exponentially with a decay length of typically 1–2 nm. It is attributed to a change in the water structure when the two surfaces approach each other. The second component is more surprising: It is very long-ranged and extends out to 100 nm in some cases. Its origin is not understood. One hypothesis is that this attraction is due to gas bubbles which form spontaneously. This is called cavitation. Estimations of the rate of cavitation, however, result in much too low values. Another hypothesis is that there are always some gas bubbles residing on hydrophobic surfaces [188]. Once these gas bubbles get into contact they fuse and cause a strong attraction due to the meniscus force. An open question remains: how these bubbles can be stable, since the reduced vapor pressure inside a bubble and the surface tension should lead to immediate collapse.

Non-DLVO forces also occur when the aqueous medium contains surfactants, which form micelles, or polyelectrolytes. A discussion of the complex interaction is, however, beyond the scope of this book. We recommend Ref. [199].

## 6.7 Steric interaction

### 6.7.1 Properties of polymers

Many dispersions are stabilised by polymers. The underlying interaction is often called the **steric force**. For the understanding of steric interactions it is necessary to know some fundamentals of polymer physics (a good introduction is the book of Strobel [190]). Here we are mainly concerned about linear polymers because these are commonly used for steric stabilization. Fortunately, in many applications we do not need to consider the detailed molecular chemical nature of the polymer such as effects of bond lengths, bond angles, rotation energy, etc. In many discussions we can use simpler models to describe the polymer.

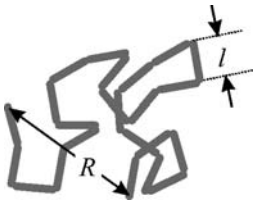



**Figure 6.14:** Picture of a linear polymer in the ideal freely jointed segments chain model.

One such model is the "ideal freely jointed segments chain" (Fig. 6.14). In this model the polymer is considered to consist of a chain of $n$ links. We call each chain link a "subunit". Each subunit has a length $l$. This parameter $l$ can correspond to the length of a monomer but it can also be shorter or longer. The angle between adjacent chain links is taken to be arbitrary. The chain forms a random coil. To characterize the size and volume of such a coil we use the mean square of the end-to-end distance $\overline{R^2}$. The square-root of this value — we call it the size of a polymer chain — is given by

$$R_0 = \sqrt{\overline{R^2}} = l\sqrt{n} \tag{6.60}$$

In light-scattering experiments, a corresponding value, the radius of gyration, is determined:

$$R_g = \frac{l\sqrt{n}}{\sqrt{6}} \tag{6.61}$$

■ **Example 6.11.** A linear polymer with a total mass of $M = 10^5$ g/mol, a monomer mass of $M_0 = 100$ g/mol, and $l = 0.5$ nm has $n = M/M_0 = 10^3$ segments assuming that each monomer corresponds to one segment. The size of the polymer is $R_0 = 0.5\,\mathrm{nm} \cdot \sqrt{10^3} = 15.8$ nm and its radius of gyration is $R_g = 6.5$ nm.

The polymer assumes this average radius only if the individual links can move freely. This is the case if we neglect the excluded volume effect due to the other segments and if we assume an "ideal" solvent. In an "ideal" solvent, the interaction between subunits is equal to the interaction of a subunit with the solvent. In a real solvent the actual radius of gyration can be larger or smaller. In a "good" solvent a repulsive force acts between the monomers. The polymer swells and $R_g$ increases. In a "bad" solvent the monomers attract each other, the polymer shrinks and $R_g$ decreases. Often a bad solvent becomes a good solvent if the temperature is increased. The temperature, at which the polymer behaves ideally, is called the theta temperature, $T_\Theta$. The ideal solvent is called the theta solvent.

### 6.7.2 Force between polymer coated surfaces

Already the ancient Egyptians knew that one can keep soot particles dispersed in water when they were incubated with gum arabicum, an exudate from the stems of acaia trees, or egg white. In this way ink was made. The reason for the stabilizing effect is the steric repulsive force cause by adsorbed polymers. In the first case these are a mixture of polysaccharide and plycoprotein, in the second case it is mainly the protein albumin. Steric stabilisation of dispersions is very important in many industrial applications. Direct quantitative measurements were

made with the surface forces apparatus [191–194] and the atomic force microscope [195–197]. Overviews are Refs. [198, 199].

The force between surfaces which are coated with polymers is mainly determined by two factors. The first one is the quality of the solvent. In good solvents the force tends to be repulsive, in bad solvents attractive. Moreover, in good solvents polymer tends to remain in solution rather than adsorbing to surfaces.

The second important factor is *how* and *how much* polymer is bound to the surface (Fig. 6.15). If the polymer is only weakly adsorbed individual molecules can still diffuse laterally. Physisorbed polymers only reach a maximal density of the order of $1/R_g^2$ and they extend typically a distance $2R_g$ into the solution. Each polymer molecules might have several binding sites and might form loops. The adsorption time plays a crucial role since polymers at surface rearrange, a process which can take hours even for liquid polymers. When the polymer is chemically bound to the surface the number of molecules on the surface — called "grafting density" $\Gamma$ — can be much higher. When the molecules are closely packed ($\Gamma \gg 1/R_g^2$) we talk about a polymer brush. In this case the steric force acts over lengths, which are substantially larger than the radius of gyration; the thickness is roughly $L_0 = nl^{5/3}\Gamma^{1/3}$, where $\Gamma$ is the grafting density in number of molecules per unit area. Polymer brushes can be made by binding a polymer to a surfaces (grafting to) or by synthesis of the polymer directly on the surface (grafting from). How such polymer layers are formed is described in Section 10.3.3.

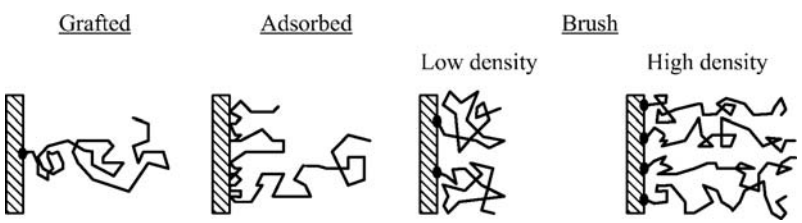



**Figure 6.15:** Structure of polymers on surfaces.

There is no simple, comprehensive theory and steric forces are complex and difficult to describe. Different components contribute to the force, and depending upon the situation, dominate the total force. The most important interaction is repulsive and of entropic origin. It is caused by the reduced configuration entropy of the polymer chains. If the thermal movement of a polymer chain at a surface is limited by the approach of another surface, then the entropy of the individual polymer chain decreases. In addition, the "concentration" of monomers in the gap increases. This leads to an increased osmotic pressure.

This repulsion was calculated by different authors [201, 202]. For a low grafting density ($\Gamma < 1/R_g^2$), the repulsive force per unit area in a good solvent and between two polymer coated surfaces is [200]

$$\Pi(x) = \frac{2k_BT\Gamma}{x} \cdot \left(\frac{2\pi^2 R_g^2}{x^2} - 1\right) \quad \text{for} \quad x \leq 3\sqrt{2}R_g \tag{6.62}$$

$$\Pi(x) = \frac{2k_BT\Gamma x}{R_g^2} \cdot e^{-\left(\frac{x}{2R_g}\right)^2} \quad \text{for} \quad x > 3\sqrt{2}R_g$$

For the other extreme, at high grafting density, the force per unit area can be approximated by [201]

$$\Pi(x) = k_B T \Gamma^{\frac{3}{2}} \cdot \left[ \left( \frac{2L_0}{x} \right)^{\frac{9}{4}} - \left( \frac{x}{2L_0} \right)^{\frac{3}{4}} \right] \tag{6.63}$$

for $x < 2L_0$. The drastic increase of the interaction with increasing grafting density is demonstrated in Fig. 6.16.

Another contribution to the steric interaction is the intersegment force. The intersegment force is caused by the direct interaction between segments of polymers with each other. This interaction depends strongly on the solvent. Below $T_\Theta$ the interaction among the monomers is stronger than the interaction of the monomers with the solvent, which results in an attractive force.

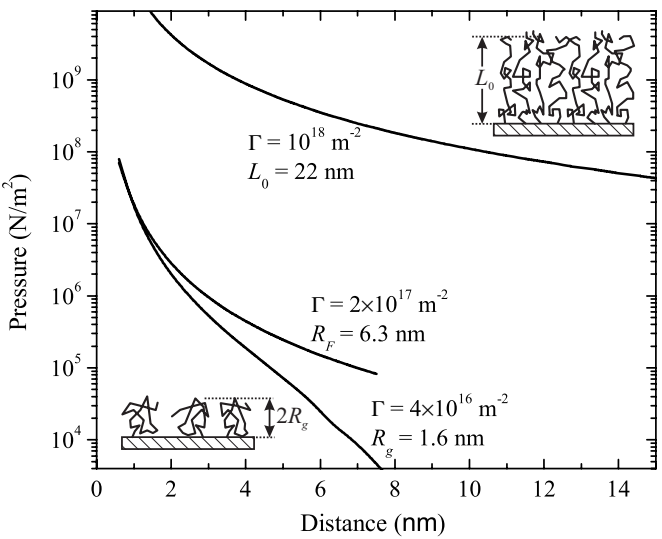



**Figure 6.16:** Repulsive steric disjoining pressure between two similar parallel plates coated with grafted polymer at grafting density $\Gamma$ in a good solvent. For the calculation we assumed a monomer length of $l = 0.4$ nm and a chain length of $n = 100$ monomers. At low grafting density ($\Gamma = 4 \times 10^{16}$ m$^{-2}$) the characteristic decay length of the force is determined by $R_g = 1.6$ nm. It was calculated using Eq. (6.63). For an intermediate grafting density ($\Gamma = 2 \times 10^{17}$ m$^{-2}$) we used [201] $\Pi = (k_B T \Gamma / R_F)\,(R_F/x)^{8/3}$, with $R_F = l n^{3/5} = 6.3$ nm. For the high grafting density ($\Gamma = 10^{18}$ m$^{-2}$) the force per unit area was calculated with Eq. (6.63) using $L_0 = n l^{5/3} \Gamma^{1/3} = 22$ nm.

Bridging forces, arising when a polymer binds to both surfaces, usually lead to an attraction at large separations. Bridging is only effective at low surface coverage. Only then do the polymer segments have a chance to find an adsorption site on the opposite surface.

■ **Example 6.12.** The steric force caused by grafted polystyrene in toluene is repulsive (Fig. 6.17). Toluene is a good solvent for polystyrene. With increasing temperature the

force becomes stronger and its range increases. This is a hint that it is caused by entropic effects.

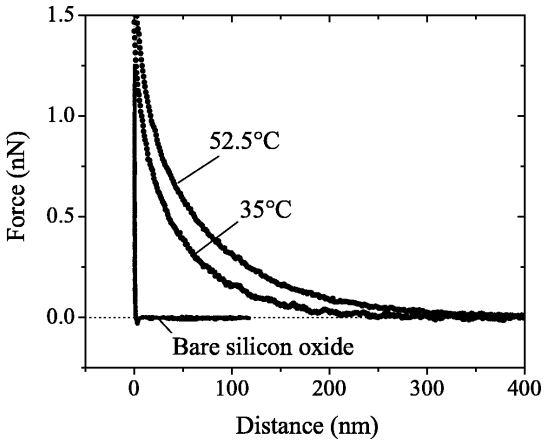



**Figure 6.17:** Steric force between an atomic force microscope tip made of silicon nitride and oxidized silicon onto which polystyrene was grafted [203]. The force was measured in toluene.

## 6.8 Spherical particles in contact

So far we have assumed that the interacting surfaces are not deformable. In reality all solids have a finite elasticity. They deform upon contact. This has important consequences for the aggregation behaviour and the adhesion of particles because the contact area is larger than one would expect from infinitely hard particles.

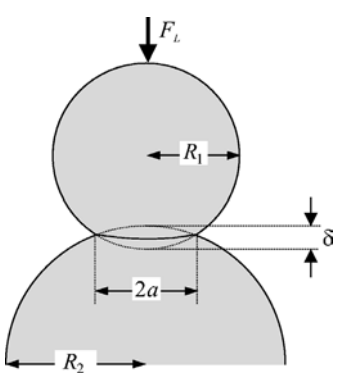



**Figure 6.18:** Schematic diagram of two spherical particles in contact.

Heinrich Hertz[8] laid the basis for the treatment of elastic solids in contact [204]. He considered two spheres with radii $R_1$ and $R_2$ in contact. The two spheres are from materials

[8] Heinrich Hertz, 1857–1894. German physicist, professor in Karlsruhe and Bonn.

with Youngs moduli $E_1$ and $E_2$ and with Poisson ratios $\nu_1$ and $\nu_2$. The two spheres are pressed together with a force $F$. This force is also called load. The load can be the gravitational force or an applied force. Hertz supposed that no surface forces act between the solids. He showed that the pressure between the spheres decreases as a quadratic function with the distance to the contact center. At the outer rim of the contact, at the radial distance $a$, the pressure is zero. If we integrate the pressure over the entire contact area we get the total force. In this way he obtained a relationship between the contact radius $a$ and the load $F$:

$$a^3 = \frac{3R^*}{4E^*} \cdot F \tag{6.64}$$

$R^*$ is the effective particle radius and $E^*$ the reduced Youngs modulus as defined by:

$$R^* = \frac{R_1 \cdot R_2}{R_1 + R_2} \qquad \text{and} \qquad \frac{1}{E^*} = \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2} \tag{6.65}$$

The contact radius at the same time determines the indentation

$$\delta = \frac{a^2}{R^*} \tag{6.66}$$

As a force-versus-distance relationship we thus get (with $a = \sqrt{\delta R^*}$ )

$$F = \frac{4}{3} \cdot E^* \cdot \sqrt{R^* \delta^3} \tag{6.67}$$

In the absence of a load $(F = 0)$ the indentation is zero and the contact radius is also zero. Since no attractive surface forces were considered there is also no adhesion in the Hertz model.

The case $a = \delta = 0$, however, never occurs and is unrealistic. The reason is that surface forces such as the van der Waals force attract the two solids even for zero load. Attractive surface forces were taken into account by Johnson, Kendall, and Robert [205]; their model is called the JKR model. They considered the following situation. If two solid surfaces come into contact, then the free surface disappears. Let us call the associated energy gain per unit area work of cohesion, $W$. The work of cohesion is correlated with an energy loss of the system. Hence, it would happen spontaneously. At the same time, work must be done for the elastic deformation of the solids. Johnson, Kendall, and Robert calculated this elastic deformation energy with the help of the Hertz theory. An enlargement of the contact area is limited by the additionally required elastic deformation energy. The real contact radius is given by

$$a^3 = \frac{3}{4} \cdot \frac{R^*}{E^*} \cdot \left( F + 3\pi W R^* + \sqrt{6\pi W R^* F + (3\pi W R^*)^2} \right) \tag{6.68}$$

For $W = 0$ we immediately obtain the Hertz result.

In the JKR model a force is necessary to separate two solids. This force is called the adhesion force. The adhesion force is

$$F_{adh} = \frac{3\pi}{2} \cdot W R^* \tag{6.69}$$

Usually the cohesion work is expressed in terms of the surface energy of the solid $\gamma_S$ by

$$W = 2\gamma_S \tag{6.70}$$

At this point, however, we need to be careful. The reason is that for solids we have to distinguish between surface tension and surface energy. The work required to form a new surface depends on how this surface is formed: plastically (as for a liquid) or elastically. In an experiment, usually both effects contribute. Therefore we should consider the surface energy obtained from adhesion experiments as an "effective" surface energy. Inserting Eq. (6.70) into Eq. (6.69) leads to

$$F_{adh} = 3\pi\gamma_S R^* \tag{6.71}$$

The adhesion force increases linearly with the particle radius. Surprisingly, it is independent of the elasticity of the materials. This is because of two opposing effects. In a hard material the deformation of the solid is small. As a result the contact area and the total attractive surface energy are small. On the other hand, the repulsive elastic component is small. Both effects compensate each other. Soft materials are strongly deformed. Thus both the attractive surface energy term and the repulsive elastic term are high.

In the JKR theory it is assumed that surface forces are active only in the contact area. In reality, surface forces are active also outside of direct contact. This is, for instance, the case for van der Waals forces. Derjaguin, Muller, and Toporov took this effect into account and developed the so-called DMT theory [206]. A consequence is that a kind of neck or meniscus forms at the contact line. As one example, the case of a hard sphere on a soft planar surface, is shown in Fig. 6.19.

Unfortunately, important results of the DMT theory cannot be expressed as convenient analytic expressions. There is, however, one simple result. For the adhesive force they obtained:

$$F_{adh} = 4\pi\gamma_S R^* \tag{6.72}$$

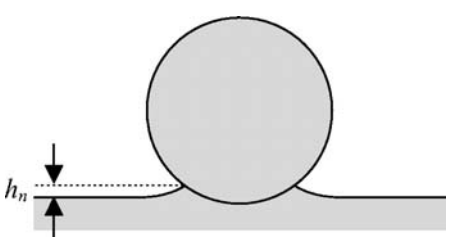



**Figure 6.19:** Neck formation according to the DMT theory.

Exact analysis shows that the two models represent two extremes of the real situation [207–209]. For large, soft solids the JKR model describes the situation realistically. For small, hard solids it is appropriate to use the DMT model . A criterion, which model is to be used, results from the height of the neck (Fig. 6.19)

$$h_n \approx \left(\frac{\gamma_S^2 R^*}{E^{*2}}\right)^{1/3} \tag{6.73}$$

If the neck height is larger than some atomic distances, the JKR model is more favorable. With shorter neck heights the DMT model is more suitable.

■ **Example 6.13.** A silicon oxide sphere with 20 $\mu$m diameter sits on a silicon oxide surface. Estimate the contact radius at negligible external force and calculate the adhesive force. $E = 5.4 \times 10^{10}$ Pa, $\nu = 0.17$, $\gamma_S = 50$ mN/m, density $\rho = 3000$ kg/m$^3$.

With $R_1 = 10$ $\mu$m and $R_2 = \infty$ the effective particle radius is $R^* = R_1 = 10$ $\mu m$.

$$\frac{1}{E^*} = 2 \cdot \frac{1 - 0.17^2}{5.4 \times 10^{10}\,\mathrm{Pa}} \Rightarrow E^* = 2.8 \times 10^{10}\,\mathrm{Pa}$$

The contact radius is estimated using the JKR model. Without external forces we have

$$\begin{aligned} a^3 &= \frac{3}{4} \cdot \frac{10^{-5}\,\mathrm{m}}{2.8 \times 10^{10}\,\mathrm{Pa}} \cdot \left(6\pi \cdot 2 \cdot 0.05\,\frac{\mathrm{N}}{\mathrm{m}} \cdot 10^{-5}\,\mathrm{m}\right) \\ &= 5.05 \times 10^{-21}\,\mathrm{m}^3 \Rightarrow a = 1.71 \times 10^{-7}\,\mathrm{m} \end{aligned}$$

The neck height is

$$h_n \approx \left(\frac{0.05^2\,\mathrm{N}^2/\mathrm{m}^2 \cdot 10^{-5}\,\mathrm{m}}{(2.8 \times 10^{10}\,\mathrm{Pa})^2}\right)^{1/3} = 3.2 \times 10^{-10}\,\mathrm{m}$$

The neck height is about as large as an atomic diameter. Therefore the DMT model is suitable and the adhesion, therefore, is

$$F_{adh} = 4\pi \cdot 0.05\,\mathrm{N/m} \cdot 10^{-5}\,\mathrm{m} = 6.3\,\mu\mathrm{N}$$

Experimentally the predictions of the theories for the behavior of spheres in contact can be measured in different ways. Adhesive forces have, for instance, been measured by a centrifuge for more than 40 years (Fig. 6.20). A significant part of the knowledge about the behaviour of powders stems from such experiments [210]. The centrifugal force, which is required to detach particles from a planar surface, is measured [211–213, 215]. Usually the detachment force of many particles is measured in a single experiment, allowing statistical evaluation of the data. This is especially useful in the case of irregularly shaped particles where the contact area and adhesion force will depend on the random orientation of the particles relative to the surface. Therefore, the centrifuge technique has been used to characterize the behavior of powders for the pharmaceutical or food industry. When tilting the surfaces to which the particles are attached, the centrifuge technique can also be used to study friction forces. There are, however, also disadvantages to this technique. One limit is that the rotational speed of the available ultra-centrifuges is limited due to the material stability of the rotor. This restricts adhesion measurements using the centrifuge method to particles larger than a few microns. Otherwise, the centrifugal force is not strong enough to detach the adhering particles from the surface. In addition, the contact time and load are difficult to vary.

Some of the disadvantages were overcome by the use of the colloidal probe technique to measure adhesion forces (review: Ref. [216]). The colloidal probe technique offers the advantage that the same particle can be used for a series of experiments and its surface can be examined afterwards. The accessible range of particle size is typically limited to a range between 1 $\mu$m and 50 $\mu$m. The tedious sample preparation, limits the number of different particles used within one study, for practical reasons. Therefore the colloidal probe and centrifugal methods complement each other.

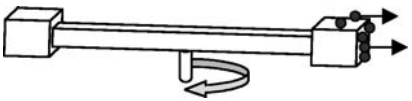


**Figure 6.20:** Schematic of the centrifuge method to determine adhesion forces of particles on surfaces. Friction forces can also be analyzed when the particles are placed on at a horizontal surface.

The JKR theory predicts correct contact radii for relative soft surfaces with effective radii larger than 100 $\mu$m. This was shown in direct force measurements by the surface forces apparatus [217, 218] or specifically designed systems. For smaller spheres it was verified using the colloidal probe technique [219].

## 6.9 Summary

- The van der Waals force includes three kinds of dipole–dipole interactions: The Keesom, the London, and the London dispersion component. Usually the dispersion interaction dominates.
- Between molecules, the van der Waals energy decreases with $1/D^6$. For macroscopic bodies the decay is less steep and it depends on the specific shape of the interacting bodies. For example, for two infinitely extended bodies separated by a gap of thickness $x$ the van der Waals energy per unit area is

$$w = -\frac{A_H}{12\pi x^2}$$

- The Hamaker constant is determined by the dielectric permittivities and the optical properties of the interacting media.
- Derjaguin's approximation allows us to calculate the force (or energy) between bodies of arbitrary shape from the force per unit area (or energy per unit area) between two planar surfaces, provided the curvature of the bodies is small compared to the typical decay lengths of the forces involved.
- Experimentally, surface forces between solids surfaces are commonly determined using the atomic force microscope or the surface force apparatus.
- In an aqueous medium, the electrostatic double-layer force is present. For distances $x$ larger then the Debye length $\lambda_D$ it decays roughly exponentially: $F \propto \exp(-x/\lambda_D)$.
- The stability of dispersions in aqueous media can often be described by the DLVO theory, which contains the double-layer repulsion and the van der Waals attraction. In some applications other effects are important, which are not considered in DLVO theory. At short range and for hydrophilic particles the hydration repulsion prevents aggregation. Hydrophobic particles, in contrast, tend to aggregate due to the hydrophobic force.

- At separations of a few molecular diameters, the solvation force due to the specific structure of the confined liquid, can be substantial. Polymers at surfaces can be used to stabilize disperions by steric interaction.

- The interaction between fine particles is often dominated by their mechanical properties such as Young's modulus. This was first considered by Hertz theory. Adhesion between spherical particles increases with the radius of the particles and is described by JKR and DMT theories.

## 6.10 Exercises

1. What is the van der Waals force per unit area for two planar, parallel layers of molecules (onemolecule thick) having surface densities (molecules per unit area) $\rho_A^\sigma$ and $\rho_B^\sigma$?

2. Verify the approximation of Eq. (6.27).

3. In some applications clays are combined with polymers, which leads to composite materials with improved thermoelastic properties. The polymer not only surrounds the clay particles but also intercalates into the layers of the clay. Thereby the polymer changes its properties because it is confined between neighboring clay layers. In one particular case we have polystyrene confined between two planar parallel layers of mica separated by 1 nm, and we are interested in the effect of the van der Waals force. To estimate this effect, please calculate the van der Waals energy per unit area $w$. Then calculate the area for which this van der Waals energy is equal to the thermal energy $k_B T$. The temperature we are interested in is 110°C, roughly the glass transition temperature of polystyrene. Do the same for silicon oxide. For simplicity, take the $\varepsilon$, $n$, and $\nu_e$ to be the same as at 20°C.

4. In coating technology, surfaces are coated to protect them from corrosion. This is usually done by a layer containing polymer, pigment, etc.. As a model system, thin polystyrene films on silicon oxide have been studied [150, 151]. Does polystyrene form a stable layer or would you expect it to dewet the surfaces assuming that van der Waals forces dominate?

5. In atomic force microscopy the tip shape is often approximated by a parabolic shape with a certain radius of curvature $R$ at the end. Calculate the van der Waals force for a parabolic tip versus distance. We only consider non-retarded contributions. Assume that the Hamaker constant $A_H$ is known.

6. Calculate the force between a conical and a parabolic object and a planar surface for an exponential force law $f = f_0 \cdot e^{-\kappa x}$.

7. Plot the contact radius versus load (up to a load of 5 $\mu$N) for two spherical powder particles of silicon oxide and a radius of $R = 2\,\mu$m using the Hertz and the JKR model ($\gamma_S = 0.05$ N/m).

8. In an aqueous electrolyte we have spherical silicon oxide particles. The dispersion is assumed to be monodisperse with a particle radius of 1 $\mu$m. Please estimate the concentration of monovalent salt at which aggregation sets in. Use the DLVO theory and assume that aggregation starts, when the energy barrier decreases below $10\,k_BT$. The surface potential is assumed to be independent of the salt concentration at $-20$ mV. Use a Hamaker constant of $0.4 \times 10^{-20}$ J.

# 7 Contact angle phenomena and wetting

This section is about the behavior of a limited amount of liquid on a solid surface. How does it spread? Which contact angle does it form? Wetting in general includes all phenomena involving contacts between three phases, of which at least two are fluid (liquid or gaseous). A typical situation is a liquid wetting a solid surface in gaseous environment. Instead of the gaseous environment we may have another imissible liquid. The wetting of a liquid by another imissible liquid is described in Chapter 13.

Wetting phenomena can be observed every day around us. Examples are rain drops on a window or dispersing a powder such as cocoa in milk. In many applications we would like to achieve complete wetting. Examples are the spreading of ink on paper, coatings and paints, the distribution of herbicides on the surface of leaves or insecticides on the epidermis of insects. In other cases we want to avoid wetting. For example, water should not wet rain-proof clothes. Road pavements should not be wetted easily by water to prevent it from penetrating into small cracks and fissures and then — in particular in winter, upon freezing — destroy the pavement. In many industrial applications like flotation or in detergency, contact angles play a decisive role. Introductions into the subject are Refs. [9, 220]. For a thorough treatment see Ref. [221].

## 7.1 Young's equation

### 7.1.1 The contact angle

Young's equation is the basis for a quantitative description of wetting phenomena. If a drop of a liquid is placed on a solid surface there are two possibilities: the liquid spreads on the surface completely (contact angle $\Theta = 0°$) or a finite contact angle is established.[1] In the second case a three-phase contact line — also called wetting line — is formed. At this line three phases are in contact: the solid, the liquid, and the vapor (Fig. 7.1). Young's equation relates the contact angle to the interfacial tensions $\gamma_S$, $\gamma_L$, and $\gamma_{SL}$ [222, 223]:

$$\gamma_L \cdot \cos\Theta = \gamma_S - \gamma_{SL} \tag{7.1}$$

If the interfacial tension of the bare solid surface is higher than that of the solid–liquid interface ($\gamma_S > \gamma_{SL}$), the right hand side of Young's equation is positive. Then $\cos\Theta$ has to be positive and the contact angle is smaller than 90°; the liquid partially wets the solid. If the solid–liquid

[1] In a strict sense the terms "complete wetting" is used for $\Theta = 0°$. "Wetting" means $\Theta < 90°$ and "non wetting" indicates that $\Theta > 90°$. De Gennes [220] distinguishes only two cases: "wetting" for $\Theta = 0°$ and "partial wetting" for $\Theta > 0°$.

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

interface is energetically less favorable than the bare solid surface ($\gamma_S < \gamma_{SL}$), the contact angle will exceed 90° because $\cos\Theta$ has to be negative.

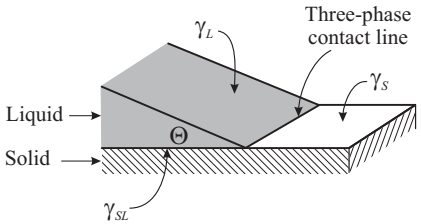



**Figure 7.1:** Rim of a liquid drop with a contact angle $\Theta$ on a solid surface.

### 7.1.2 Derivation

We derive Young's equation for a typical example: that of a circular drop on a planar solid surface. Other geometries can be treated in an analogous way. To derive Young' s equation we consider the change of Gibbs free energy $dG$ when the drop spreads an infinitesimal amount. While spreading, the radius of the contact zone increases from $a$ to $a' = a + da$ (Fig. 7.2). Assuming constant volume, its height decreases from $h$ to $h' = h + dh$ ($dh$ is negative). If the change in Gibbs free energy is negative, the process is going to occur spontaneously. For positive $dG$ the drop will contract. In equilibrium, which is the energetically most favorable situation, we have $dG = 0$.

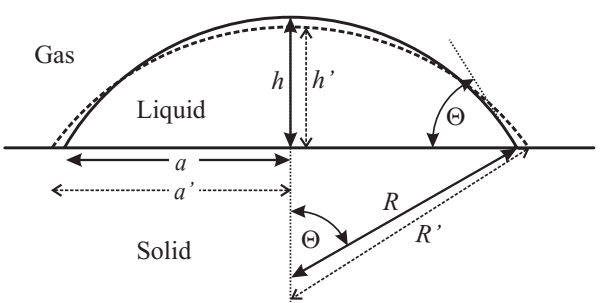



**Figure 7.2:** Liquid drop with circular contact area on a planar solid surface.

On one hand we assume that the drop is so large that a contribution of the Laplace pressure does not lead to a significant change in Gibbs free energy; for very small drops we have to introduce a correction. On the other hand, we assume that the drop is small enough, so that gravitation is negligible. As a consequence its shape is that of a spherical cap. The result is also valid for large drops but for negligible gravitation the mathematical treatment is simpler.

When the drop spreads, the free solid surface is changed to a solid–liquid interface. The change in area $dA_{SL} = 2\pi a\, da$ leads to a change in surface energy $(\gamma_{SL} - \gamma_S)\, dA_{SL}$. In addition, the surface area of the liquid–gas interface changes. Elementary geometry tells us

that the surface of a spherical cap is

$$A_L = \pi \left(a^2 + h^2\right) \tag{7.2}$$

A small change in the contact radius $a$, leads to a change in the liquid surface area of

$$dA_L = \frac{\partial A_L}{\partial a} da + \frac{\partial A_L}{\partial h} dh = 2\pi a\, da + 2\pi h\, dh \tag{7.3}$$

Unfortunately, the change in surface area depends on two variables: $a$ and $h$ (we do not have to consider a change in $\Theta$, it is of second order). However, these two variables are not independent because the volume of the drop is constant. The volume of a spherical cap is

$$V = \frac{\pi}{6}\left(3a^2 h + h^3\right) \tag{7.4}$$

A small change of this volume is

$$dV = \frac{\partial V}{\partial a}\, da + \frac{\partial V}{\partial h}\, dh = \pi a h\, da + \frac{\pi}{2}\left(a^2 + h^2\right)\, dh \tag{7.5}$$

Since the volume is assumed to be constant ($dV = 0$) we have

$$-\pi a h\, da = \frac{\pi}{2}\left(a^2 + h^2\right)\, dh \Rightarrow \frac{dh}{da} = -\frac{2ah}{a^2 + h^2} \tag{7.6}$$

Using Pythagoras' law, $R^2 = a^2 + (R - h)^2 \Rightarrow a^2 = 2Rh - h^2$, we get

$$\frac{dh}{da} = -\frac{2ah}{2Rh - h^2 + h^2} = -\frac{a}{R} \tag{7.7}$$

and

$$\begin{aligned} dA_L &= 2\pi a\, da - 2\pi h \frac{a}{R}\, da = 2\pi a \cdot \left(1 - \frac{h}{R}\right)\, da \\ &= 2\pi a\, \frac{R - h}{R}\, da = 2\pi a\, \cos\Theta\, da \end{aligned} \tag{7.8}$$

Now we can write the total change in the Gibbs free energy as

$$\begin{aligned} dG &= \left(\gamma_{SL} - \gamma_S\right)\, dA_{SL} + \gamma_L\, dA_L \\ &= 2\pi a\, \left(\gamma_{SL} - \gamma_S\right)\, da + 2\pi a\, \gamma_L \cdot \cos\Theta\, da \end{aligned} \tag{7.9}$$

At equilibrium $dG/da = 0$ which directly leads to Young's equation .

One should keep in mind that Young' s equation is only valid in thermodynamic equilibrium, hence in the presence of a saturated vapor of the liquid. In most practical applications this is not the case.

Young's equation is also valid if we replace the gas by a second, immiscible liquid. The derivation would be the same, we only have to replace $\gamma_L$ and $\gamma_{SL}$ by the appropriate interfacial tensions. For example, we could determine the contact angle of a water drop on a solid surface under oil. Instead of having a gas saturated with the vapor, we require to have a second liquid saturated with dissolved molecules of the first liquid.

### 7.1.3 The line tension

Spreading is usually accompanied by a change of the length of the wetting line. For example, if our drop with a circular contact area spreads, the length of the three phase contact line increases by $2\pi a\, da$. Just as the formation of a new surface area, the creation of a wetting line also requires energy. This energy per unit length is called line tension $\kappa$. For drops much smaller than 1 mm the line tension has to be taken into account and a term has to be added to Young's equation [224]:

$$\gamma_L \cdot \cos\Theta = \gamma_S - \gamma_{SL} - \frac{\kappa}{a} \tag{7.10}$$

Here, $a$ is the radius of curvature of the three-phase contact line. For a drop with circular contact area it is the contact radius.

■ **Example 7.1.** Estimate the line tension for cyclohexane liquid from its heat of vaporization $\Delta_{vap}U = 30.5$ kJ/mol at 25°C. Then the density is $\rho = 773$ kg/m$^3$, its molecular weight is $M = 84.2$ g/mol.

As in example 2.3, we picture the liquid as being composed of cubic molecules. The size of each cube was calculated from the density of cyclohexane to be $a = 0.565$ nm. In the bulk each molecule is supposed to directly interact with 6 neighbors. The energy per bond is thus $\Delta_{vap}U/6N_A$. At the rim two bonds less can be formed and the energy loss per molecule is $2\Delta_{vap}U/6N_A$. Thus the energy difference per unit length is

$$\kappa = \frac{\Delta_{vap}U}{3N_A\, a} = \frac{30500\ \mathrm{Jmol}^{-1}}{3 \cdot 6.02 \times 10^{23}\,\mathrm{mol}^{-1} \cdot 0.565 \times 10^{-9}\,\mathrm{m}} = 2.98 \times 10^{-11}\ \frac{\mathrm{J}}{\mathrm{m}}$$

The result leads to the right order of magnitude. For water, the same calculation results in an estimated line tension of $7.4 \times 10^{-11}\ \mathrm{J/m}$.

Typical line tensions are in the order of $10^{-10}$ N [226]. In some cases significantly higher *effective* line tensions have been determined [225]. Calculations of the line tension, which are based on the analysis of surface forces, are reviewed in Ref. [227].

### 7.1.4 Complete wetting

We can rearrange Young' s equation and get

$$\cos\Theta = \frac{\gamma_S - \gamma_{SL}}{\gamma_L} \tag{7.11}$$

The cosine cannot exceed one. Then we might ask: What happens, if $\gamma_S - \gamma_{SL} - \gamma_L > 0$ or $\gamma_S - \gamma_{SL}$ is higher than $\gamma_L$? Does this not violate Young's equation? No, it does not because in thermodynamic equilibrium $\gamma_S - \gamma_{SL} - \gamma_L$ can never become positive. This is easy to see. If we could create a situation with $\gamma_S > \gamma_{SL} + \gamma_L$, then the Gibbs free energy of the system could decrease by forming a continuous liquid film on the solid surface. Vapor would condense onto the solid until such a film is formed and the free solid surface would be replaced to a solid–liquid interface plus a liquid surface.

At first glance the case $\gamma_S = \gamma_{SL} + \gamma_L$ looks like an exception. That, however, is not the case. In equilibrium, in particular in the presence of saturated vapor, $\gamma_S$ can never become larger than $\gamma_L + \gamma_{SL}$. Hence, if the system is in equilibrium and we have complete wetting then $\gamma_S = \gamma_{SL} + \gamma_L$.

In most practical applications the system is not in perfect thermodynamic equilibrium. In non-equilibrium situations the so-called spreading coefficient (see also Eq. 13.18)

$$S = \gamma_S - \gamma_{SL} - \gamma_L \tag{7.12}$$

can indeed be positive. It is a measure of how strongly a liquid can spread on a surface. For $S < 0$ a finite contact angle is formed.

## 7.2 Important wetting geometries

### 7.2.1 Capillary rise

An important application of Young's equation and a procedure for measuring the contact angle at the same time, is the rise of a liquid in a capillary tube. If a capillary is lowered into a liquid, the liquid often rises in the capillary until a certain height is reached (Fig. 7.3 top left). For a capillary with a circular cross-section of radius $r_C$, this height is given by

$$h = \frac{2\gamma_L \cdot \cos\Theta}{r_C g \rho} \tag{7.13}$$

Here, $g$ is the acceleration of free fall and $\rho$ is the density of the liquid.

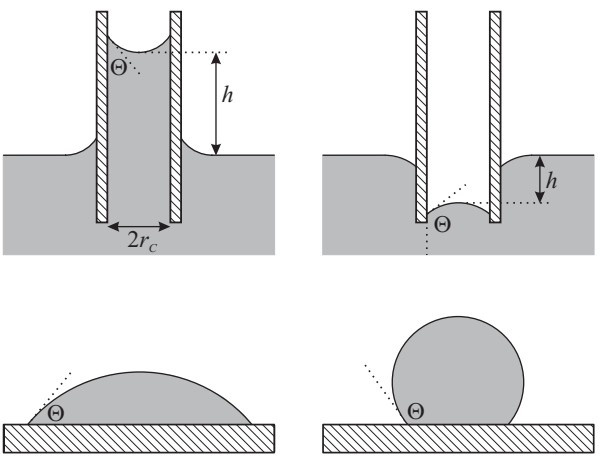



**Figure 7.3:** Rise of a liquid in a partially wetted (left) and capillary depression in a nonwetted (right) capillary. Below, the same wetting situations are shown for a drop on a planar solid surface.

■ **Example 7.2.** Water in trees rises in capillaries which are called xylem. They are typically 5–170 $\mu$m in radius and are completely wetted ($\Theta = 0$). What is the maximum height water can rise in such capillaries? To calculate an upper limit we use a capillary radius of 5 $\mu$m. Then

$$h = \frac{2 \cdot 0.072\ \mathrm{Nm}^{-1}}{5 \times 10^{-6}\mathrm{m} \cdot 9.81\ \mathrm{ms}^{-2} \cdot 997\ \mathrm{kgm}^{-3}} = 2.9\ \mathrm{m} \tag{7.14}$$

This is certainly not enough to bring water from the soil to the top of a tree. In fact, other effects contribute significantly. For details see Ref. [228].

To derive Eq. (7.13) we consider the change of Gibbs free energy upon an infinitesimal rise of the liquid $dh$. This is simple because the shape of the liquid–vapor interface does not change. The change in Gibbs free energy is:

$$dG = -2\pi r_C \cdot dh \cdot (\gamma_S - \gamma_{SL}) + \pi r_C^2 \rho g h \cdot dh \tag{7.15}$$

The first term represents the surface work. This is actually driving the process: the liquid rises because high-energy solid surface is exchanged for a low-energy solid–liquid interface. The second term corresponds to the work required to lift the liquid with a weight $\pi r_C^2 h \rho$ by a step $dh$ against gravity. In equilibrium, the gravitation term is equal to the contribution of the interfacial tensions and we have

$$\frac{dG}{dh} = -2\pi r_C \cdot (\gamma_S - \gamma_{SL}) + \pi r_C^2 \rho g h = 0 \Rightarrow 2\,(\gamma_S - \gamma_{SL}) = r_C \rho g h \tag{7.16}$$

After replacing $\gamma_S - \gamma_{SL}$ by $\gamma_L \cos\Theta$ we immediately get Eq. (7.13).

Not only does a capillary rise occur but, depending on the interfacial tensions, liquid can also be expelled from a capillary. A liquid rises for partially wetted surfaces ($\Theta < 90°$). If the liquid does not wet the inner surface of the capillary and the contact angle is higher than 90° the liquid is pressed out of the capillary. For this reason it is very difficult to get water into polymeric capillaries, as long as the capillary is hydrophobic. This can achieved only by applying an external pressure.

### 7.2.2 Particles in the liquid–gas interface

Small particles bind to liquid–gas interfaces if the contact angle is not zero. As we will see later this is of fundamental importance in applications such as flotation or the stabilization of emulsions. For simplicity we start by considering small, spherical particles. "Small" refers to particles for which we can neglect gravitational effects and buoyancy. Practically, this is valid for particles with diameters in the range of up to $\approx 100$ $\mu$m.

For $\Theta > 0$ a particle is stable at the liquid surface (Fig. 7.4) [229]. Its position in the interface is simply determined by the fact that, in equilibrium, the liquid surface is unperturbed. A planar liquid surface will also be planar with the adsorbed particle. This is easy to understand. If the liquid surface is curved, the capillary force due to surface tension would result in a force in the normal direction. In the absence of external forces, this capillary force would draw the particle to a new position, until the curvature is zero again. If the particle is significantly larger

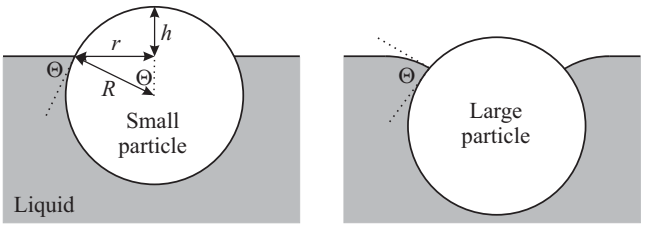



**Figure 7.4:** Small spherical particle in a liquid–gas interface for which gravitation is negligible (left). The shape of the interface remains unchanged by the presence of the particle. For larger particles (right) the liquid interface is deformed and a net capillary force stabilizes the particle and prevents it from sinking.

than 100 $\mu$m, gravitation is pulling it down. The liquid surface becomes curved so that the capillary force compensates for the gravitational force.

For many applications the work required to remove a particle from the gas–liquid interface is important. For example, in flotation (Section 7.6.1) mineral particles adhere to gas bubbles and we would like them to remain in the interface rather than going back into the bulk liquid. This work is equal to the change in the Gibbs free energy. We calculate the change in Gibbs free energy upon moving a small particle from its equilibrium position in the interface into a bulk liquid, where it is completely immersed. Therefore, we first need to know the surface area of the particle which is exposed to the gas phase. It is (see also Eq. 7.2)

$$\pi\left(r^2+h^2\right)=\pi R^2\cdot\left(\sin^2\Theta+1-2\cos\Theta+\cos^2\Theta\right)=2\pi R^2\cdot(1-\cos\Theta) \quad (7.17)$$

with $r = R\sin\Theta$ and $h = R - R\cos\Theta$. When moving the particle from the interface into liquid, the solid–liquid and the liquid–vapor areas increase while solid–gas interfacial area decreases. This results in a change in Gibbs free energy of

$$\Delta G = 2\pi R^2\cdot(1-\cos\Theta)\cdot(\gamma_{SL}-\gamma_S)+\pi R^2\sin^2\Theta\cdot\gamma_L \quad (7.18)$$

With Young's equation we get

$$\begin{aligned}\Delta G &= -2\pi R^2\cdot(1-\cos\Theta)\cdot\gamma_L\cos\Theta+\pi R^2\sin^2\Theta\cdot\gamma_L\\ &= \pi R^2\gamma_L\cdot\left(\sin^2\Theta-2\cos\Theta+2\cos^2\Theta\right)\\ &= \pi R^2\gamma_L\cdot\left(\cos^2\Theta-2\cos\Theta+1\right)\\ &= \pi R^2\gamma_L\cdot(\cos\Theta-1)^2 \quad (7.19)\end{aligned}$$

■ **Example 7.3.** To remove a hydrophobic microsphere ($\Theta = 90°$) of 5 $\mu$m radius from a water surface ($\gamma_L = 0.072\,\mathrm{Jm^{-2}}$) work of $5.6\times10^{-12}$ J is required. This is roughly $10^9\ k_BT$.

We can estimate the force necessary to remove the particle from the gas–liquid interface from $F \approx \Delta G/R = 1.1\ \mu$N. Compared to the gravitational force, which is of the order of $10^{-11}$ N, this is strong. We conclude that under normal conditions the particle will remain in the interface.

### 7.2.3 Network of fibres

For many applications the wetting behavior of a network of fibres is important. An example is the water repellent ability of clothing. As a simple model we consider a bundle of parallel cylinders which are separated by a certain spacing. This spacing is assumed to be small compared to the capillary constant, so that the shape of the liquid surface is assumed to be determined only by the Laplace equation. Practically, this leads to cylindrical liquid surfaces.

For small external pressures such a network will not allow a liquid to pass, unless the contact angle is zero (Fig. 7.5) [230]. The liquid forms the contact angle with the solid which determines how far it will penetrate into the interfibrous spacing. If an external pressure such as hydrostatic pressure is applied, water moves further into the interfibrous spacing. It penetrates the network only if the pressure is so high that, for a given spacing, the liquid surface becomes unstable and a drop is formed.

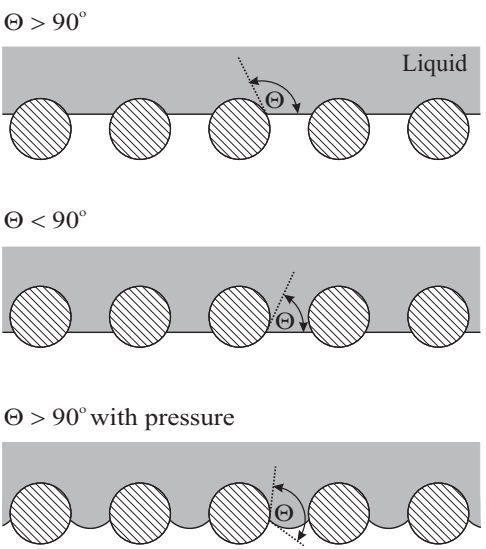



**Figure 7.5:** Parallel cylindrical fibres with a liquid on top. In the first two cases no external pressure, such as hydrostatic pressure, is applied. For this reason the liquid surface is planar (in equilibrium). The contact angle is larger than 90°, in the first case and smaller than 90° but significantly higher than 0°, in the second case. In the third case a hydrostatic pressure is applied.

■ **Example 7.4.** Effect of oil on aquatic birds. Aquatic birds keep their feathers hydrophobic with wax. For them it is essential that the feathers repel water. The air entrapped in and under the feathers provides a good heat insulation. It keeps the birds afloat and light so that they are able to fly. Due to low surface tension, oil has a strong tendency to wet all kinds of solid surfaces. It also wets the surfaces of feathers which then deprives the birds of all the essentials described. In addition, heavy oil is sticky, destroys the feather structure, and prevents the birds from moving freely.

## 7.3 Measurement of the contact angle

### 7.3.1 Experimental methods

The most common method of measuring the contact angle is to observe a **sessile drop** with a telescope or microscope. A light source is positioned behind the drop, which then appears dark. The contact angle is either determined directly with a goniometer or the image is recorded by a video system and the contour is fitted by a computer by the Laplace equation. The same procedure can also be used to determine the surface tension of the liquid. For small drops, where hydrostatic effects are negligible, we can also calculate the contact angle from the height $h$ of the drop (or from other easily measurable parameters) [231]. As predicted by the Laplace equation small drops have a circular cross-section (Fig. 7.6). From the measured height and contact radius $a$ of the drop we can calculate the contact angle by

$$\tan\left(\frac{\Theta}{2}\right) = \frac{h}{a} \tag{7.20}$$

Alternatively, we can measure the contact angle at the edge of a bubble. This method is called **captive** or **sessile bubble**. In this case a bubble is positioned usually at the top of a cell which is otherwise filled with liquid. The method is less sensitive to pollution of the interface. In addition, the vapor phase is automatically saturated.

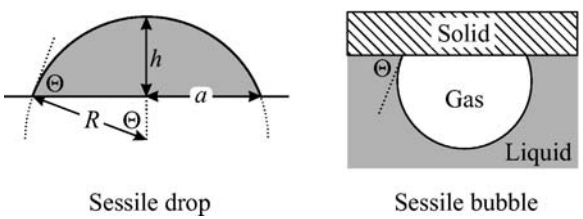



**Figure 7.6:** Sessile drop and sessile bubble on a planar surface.

A widely used technique is the **Wilhelmy plate method** introduced in Section 2.4. If the contact angle is larger than zero, the force, with which the plate is pulled into the liquid, is $2\gamma_L l \cos\Theta$. Here, $l$ is the width of the plate.

Contact angles can have quite different values. Examples of contact angles at room temperature are given in Table 7.1.

In many applications, powders come into contact with a liquid and we would like to quantify their wetting behavior . The usual way to do this is by the **capillary rise method** [233, 234]. In a capillary rise measurement the powder is pressed into a tube of typically 1 cm diameter (Fig. 7.7). This porous material is then treated as a bundle of thin capillaries with a certain effective radius $r_C^{eff}$ [235–237]. In order to measure this effective radius, first a completely wetting liquid is used. Either the speed of the liquid rise is measured (this technique is sometimes referred to as the capillary penetration technique [238]) or the pressure required to keep the liquid out of the porous material, is determined. This backpressure is equal to the

| Liquid | Solid | $\Theta_{adv}$ |
|---|---|---|
| Water | Paraffin | 110° |
| | Teflon | 98–112° |
| | Polystyrene | 94° |
| | Polypropylene | 108° |
| | Polyethylene | 88–103° |
| | Graphite | 86° |
| | Gold* | 0° |
| $n$–decane | Teflon | 32–40° |
| $n$-octane | Teflon | 26–30° |
| $n$-propanol | Teflon | 43° |
| | Paraffin | 22° |
| | Polyethylene | 7° |
| Mercury | Teflon | 150° |
| | Glass | 128–148° |

**Table 7.1:** Contact angles (advancing) of different liquids on various solids at 25°C (from Refs. [1,232] and own measurements).
* If you measure the contact angle on freshly prepared gold, you probably find $\Theta = 20° \ldots 70°$. This is due to contamination with organic substances. For clean gold the contact angle is zero.

Laplace pressure of the liquid, given by

$$\Delta P_{ref} = \frac{2\gamma_L^{ref}}{r_C^{eff}} \tag{7.21}$$

Knowing $\gamma_L^{ref}$ we obtain $r_C^{eff}$. Then we measure the Laplace pressure with the liquid of interest. It is

$$\Delta P = \frac{2\gamma_L \cos\Theta}{r_C^{eff}} \tag{7.22}$$

Comparing the two pressures directly leads to the contact angle. One limitation of the capillary rise method is that it averages over many particles and the actual size distribution remains unknown. In addition, it relies on the assumption that a powder can be treated as a bundle of capillaries and depends on the specific model applied [239, 240].

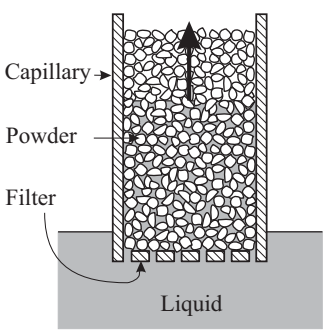



**Figure 7.7:** Capillary rise method to quantify the wetting properties of powders or porous materials.

Recently it became possible to determine the contact angle of individual spherical particles by using a technique called microsphere tensiometry [241]. It is closely related to the colloidal

probe technique (see Section 6.4). In this technique the particle is placed into the gas–liquid interface. The equilibrium position is precisely determined. Knowing the radius of the particle from electron microscopy, the contact angle can easily be calculated from the geometry.

### 7.3.2 Hysteresis in contact angle measurements

So far we have dealt with ideal surfaces. On real surfaces we have to face the difficulty of hysteresis. Let us illustrate this for the most common situation, that of a sessile drop on a planar surface. If we measure the contact angle while the volume of the drop is increasing — practically this is done just before the wetting line starts to advance — we get the so-called advancing contact angle $\Theta_{adv}$. If we afterwards decrease the volume of the drop and determine the contact angle just before the wetting line is receding, we measure the so-called receding contact angle $\Theta_{rec}$. Usually $\Theta_{adv}$ is significantly higher than $\Theta_{rec}$. The difference $\Theta_{adv} - \Theta_{rec}$ is called contact angle hysteresis. It is typically 5–20° but can also be significantly higher. Hysteresis complicates the mathematical treatment considerably because we do not know which situation refers to thermodynamic equilibrium.

What are the possible reasons for hysteresis? Various explanations for contact angle hysteresis have been discussed. The contribution of each factor depends on the particular situation encountered. Possible causes of hysteresis are:

- Surface roughness [242–244]. Although many surfaces appear flat and homogeneous to the naked eye, at the submicroscopic scale they are usually rough. This can lead to hysteresis as illustrated for a drop with a contact angle of 90°, in Fig. 7.8. Here, the roughness is illustrated by a microscopically small single spherical protrusion. If the three-phase contact line advances from left to right (position A) it will at some point come into contact with the protrusion. There it jumps immediately to a position, where it again assumes the contact angle of 90° (B). In our case this is at the top of the protrusion. Then the drop is hindered from spreading further in order to keep locally a contact angle of 90°. Macroscopically the three-phase contact line sticks at the protrusion. The apparent contact angle is now much larger than 90° (C). Only after overcoming the protrusion will it continue to spread spread until it reaches the next protrusion. A rough surface can be pictured as a collection of densely packed protrusions. They might not be as steep as in our example but the effect is qualitatively similar. When the liquid recedes, the same effect occurs and we observe a hysteresis.

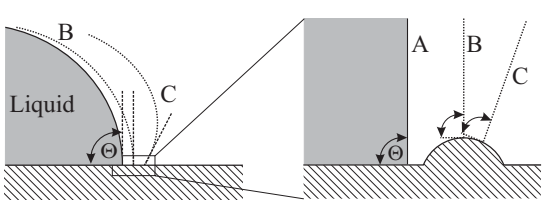



**Figure 7.8:** A drop advancing over a solid surface with a microscopic protrusion. The left side shows the optically visible situation. On the right, microscopic details are schematically illustrated.

- Most surfaces are not perfectly homogeneous but expose chemically or structurally different regions. Such heterogeneity or contamination of the solid surface leads to line pinning and thus to hysteresis [245–249] just as described for roughness. Upon advancing, the three-phase contact line is pinned by relatively lyophobic ("liquid hating") regions, when receding the three-phase contact line is pinned by lyophilic ("liquid loving") regions.

- Dissolved substances often adsorb at the three-phase contact line. Advancing or receding of the liquid is hindered by the deposited substances.

- At the three-phase contact line the surface tension exerts strong forces on the surface. For instance, if we consider a water drop on a polymer surface, typical contact angles are $90°$. The surface tension pulls upwards on the solid surface. If we estimate the wetting line to have a width of $\delta = 10$ nm, the force $F$ per unit length $l$ can be related to the effective pressure exerted on the solid surface:

$$\gamma = \frac{F}{l} = \frac{P \cdot l \cdot \delta}{l} = P \cdot \delta \tag{7.23}$$

  With $\gamma = 0.072\ \mathrm{Nm}^{-1}$ and $\delta = 10$ nm the effective pressure is of the order of $P = 72 \times 10^5$ Pa. Such a high pressure can change the surface structure, cause mechanical deformation at the moving wetting line [250], and can lead to contact angle hysteresis [251–253], especially on soft polymer surfaces.

- Adsorption and desorption of liquid molecules at the spreading or receding liquid are accompanied by the dissipation of energy and are thus one source of hysteresis [254,255]. At this point we would like to point out the similarity between contact angle and adhesion experiments. Adhesion is dominated by the solid–solid attraction, while contact angles reflect the solid–liquid attraction.

The influence of roughness, inhomogeneity, line tension, energy dissipation etc. is not always direct but indirect and they might be effective in combination [256]. Heterogeneity or surface roughness can, for instance, lead to line pinning at certain points. Microscopically, the wetting line looks like a random zigzag [257]. While the liquid spreads or retreats it constantly changes its shape. Since in most cases the formation of a wetting line requires energy, this process leads to energy dissipation.

A very important factor is the length scale of surface structures and their shape. This becomes obvious when looking at the spreading of a liquid on surfaces with a micropatterned heterogeneity [246, 258, 259].

### 7.3.3 Surface roughness and heterogeneity

In the preceding section, we discussed the fact that roughness and heterogeneity of surfaces can lead to contact angle hysteresis. In addition, it influences the mean contact angle. An important aspect is the length scale involved. Here, we consider surfaces which are rough and heterogeneous on a scale above the molecular size but below a scale, which is easily observable with optical techniques. This is usually 1 nm to 1 $\mu$m.

For not too rough surfaces[2] we can describe the effect of surface roughness by the so-called Wenzel equation [260]:

$$\cos\Theta_{app} = R_{rough} \cdot \cos\Theta \tag{7.24}$$

Here, $\Theta_{app}$ is the apparent contact angle which we observe by eye or with an optical microscope. $R_{rough}$ is the ratio between the actual and projected surface area. Since $R_{rough}$ is always larger or equal to one ($R_{rough} \geq 1$) surface roughness decreases the contact angle for $\Theta < 90°$, while for poorly wetted surfaces ($\Theta > 90°$) the contact angle increases. If a molecularly hydrophobic surface is rough, the appearance is that of an even more hydrophobic surface. If a hydrophilic surface is roughened it becomes more hydrophilic. For an instructive discussion see Ref. [261].

A striking example is the so-called Lotus effect [262]: On some plant leaves the contact angle of water is very large and water drops roll off, picking up dirt on their way and thus cleaning the surface. The surface of these leaves are hydrophobic and extremely rough. They can be viewed more like a surface with many hydrophobic spikes (Fig. 7.9). A water drop placed on top only comes into contact with the ends of the spikes. Air is trapped below the drop so that the contact angle increases. If there were only air, the contact angle would reach a maximum value of $180°$. For some technical applications this is used to create surfaces which are self-cleaning. Dirt particles are easily washed away by rain or water. Using this effect, super-water-repellent (also called superhydrophobic) surfaces, are made [263–265].

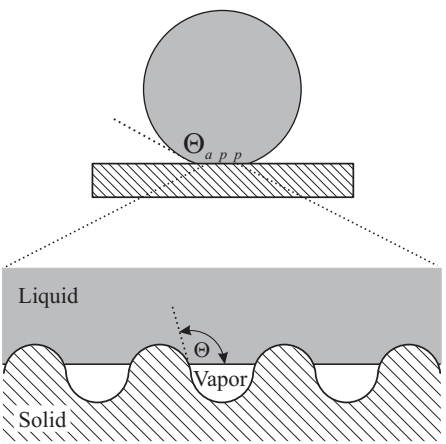



**Figure 7.9:** Water drop on a superhydrophobic surface showing a high apparent contact angle $\Theta_{app}$. The combined effect of hydrophobicity and roughness on the right length scale, causes the Lotus effect.

Most solid surface are also chemically inhomogeneous. Cassie considered a smooth but chemically patchwise heterogeneous surface. If there are two different kinds of region with

[2] "Not too rough" here refers to surfaces which look smooth and glossy. Such an optical appearance implies, that the roughness is significantly below the wavelength of light.

contact angles $\Theta_1$ and $\Theta_2$, which occupy the surface ratios $f_1$ and $f_2$ the apparent average contact angle is [266]

$$\cos\Theta_{app} = f_1 \cos\Theta_1 + f_2 \cos\Theta_2 \tag{7.25}$$

■ **Example 7.5.** Drelich et al. [246] measured contact angles of water drops on patterned surfaces. They created parallel stripes of hexadecanethiol ($HS(CH_2)_{16}CH_3$) and diundecane disulfide carboxylic acid ($S[(CH_2)_{11}COOH]_2$) on gold, which were 2.5 and 3.0 $\mu$m wide, respectively. Pure hexadecanthiol monolayers (see Section 10.3.1) are hydrophobic and showed an advancing contact angle of 107.8° at pH 7.0. The disulfides were more hydrophilic (advancing contact angle 50.1°) because the carboxylic acid is exposed on the surface. For the micropatterned surface, a contact angle of

$$\cos\Theta_{app} = \frac{2.5}{2.5+3.0}\cos 107.8^\circ + \frac{3.0}{2.5+3.0}\cos 50.1^\circ = 0.211 \Rightarrow$$

$$\Theta_{app} = 77.8^\circ$$

is calculated using the Cassie equation (7.25). Experimentally, they observed an advancing contact angle of $77^\circ \pm 3^\circ$ parallel to the stripes, which is in good agreement. Perpendicular to the stripes the experimental contact angle deviated significantly from the calculated value and the authors suggest to use a modified version of the Cassie equation.

## 7.4 Theoretical aspects of contact angle phenomena

Young's equation is extensively used to quantify wetting. Unfortunately, it is not well founded and it has not really been verified empirically. A major problem is to determine the surface tensions of the solid $\gamma_S$ and $\gamma_{SL}$ [46,267]. In most applications the surface tension and energy depends on the way the surface has been created and in this sense is not a real thermodynamic quantity. Solid surfaces can be formed plastically or elastically. Depending on the method of formation, the surface tension or energy can be different. Another severe problem is that, experimentally, only advancing or receding contact angles can be determined. They only give a range for the equilibrium contact angle.

The parameters $\gamma_S$, $\gamma_L$, $\gamma_{SL}$, and $\Theta$ describe the situation on the macroscopic scale, at a distance of more than 0.1–1 $\mu$m from the wetting line. Close to the wetting line, in the core region, the situation is more complicated. Here, surface forces can change the shape of the drop and we have to distinguish between the microscopic contact angle (in the region a few nm from the wetting line) and the macroscopic contact angle (which can be observed by optical methods, including optical microscopy). It was one of the great achievements of Young to realize that it is possible to calculate the macroscopic contact angle from the interfacial tensions without knowing the precise shape in the core region.

Within the core region the profile of a drop is modified by surface forces, such as long-range van der Waals and electrostatic double-layer forces [220]. These forces affect the profile in a range of 1–100 nm. They can cause a difference between the microscopic contact angle and the macroscopic one (which enters into Young's equation) [268, 269]. If the liquid is attracted by the solid surface and this attraction is stronger than the attraction between the

liquid molecules, we have formally a repulsive disjoining pressure between the solid–liquid and the liquid–gas interfaces. It tends to thicken the liquid film (Fig. 7.10 A). In contrast, if the liquid molecules attract each other more strongly than they are attracted by the solid we have formally an attractive adjoining pressure between the solid–liquid and the liquid–gas interfaces (Fig. 7.10 B).

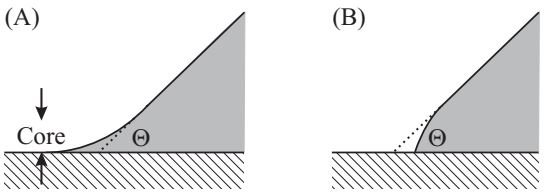



**Figure 7.10:** Microscopic view of the three-phase contact region for a repulsive force between the solid–liquid and liquid–gas interface (A) and an attractive force between the two interfaces (B).

In many applications it is possible to determine $\gamma_L$ and, with more difficulty, $\gamma_S$, but not $\gamma_{SL}$. Therefore, it would be helpful to express $\gamma_{SL}$ through $\gamma_L$ and $\gamma_S$. Since $\gamma_{SL}$, $\gamma_L$, and $\gamma_S$ are independent parameters we can only hope to find an approximate expression and we have to use additional information. Girifalco, Good, and Fowkes considered solids and liquids, in which the molecules are held together by van der Waals forces [270, 271]. Then, in a thought experiment they separated two materials at the interface (Fig. 7.11). The required work of adhesion per cross-sectional area is $w = \gamma_1 + \gamma_2 - \gamma_{12}$. Two new surfaces are formed while the interfacial area disappears. Rearrangement leads to

$$\gamma_{12} = \gamma_1 + \gamma_2 - w \tag{7.26}$$

Now we recall Section 6.2.3. The work of separating two solids to an infinite distance against the van der Waals attraction, is given by Eq. (6.29)

$$w = A_{12}/12\pi D_0^2 \tag{7.27}$$

Here, $D_0$ is a typical interatomic spacing and $A_{12}$ is the Hamaker constant for the interaction of material 1 with material 2 across a vacuum or gas. According to Eq. (6.28) the "mixed" Hamaker constant can be expressed by

$$A_{12} \approx \sqrt{A_{11} \cdot A_{22}} \tag{7.28}$$

The Hamaker constants of the single materials are related to the surface tensions

$$\gamma_1 = A_{11}/24\pi D_0^2 \qquad \text{and} \qquad \gamma_2 = A_{22}/24\pi D_0^2 \tag{7.29}$$

Substituting all these results into Eq. (7.26) finally leads to

$$\gamma_{12} = \gamma_1 + \gamma_2 - 2\sqrt{\gamma_1\gamma_2} \tag{7.30}$$

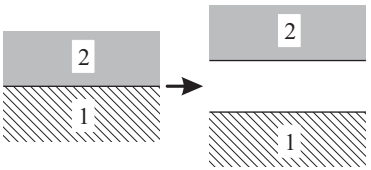



**Figure 7.11:** Two materials are separated at their interface.

The model of Girifalco, Good, and Fowkes has been extended to other interactions. For example, if we assume that the surface energies are the sum of van der Waals (dispersive) and polar interactions, one often uses the equation [272]

$$\gamma_{SL} \approx \gamma_S + \gamma_L - 2\sqrt{\gamma_S^d \gamma_L^d} - 2\sqrt{\gamma_S^p \gamma_L^p} \tag{7.31}$$

with $\gamma_S = \gamma_S^d + \gamma_S^p$ and $\gamma_L = \gamma_L^d + \gamma_L^p$. The superscripts "$d$" and "$p$" indicate dispersive and polar interactions. For a discussion see Refs. [273, 274].

## 7.5 Dynamics of wetting and dewetting

### 7.5.1 Wetting

In dynamic wetting a liquid displaces another fluid, often air, from a solid surface. We can distinguish forced and spontaneous wetting. In forced wetting, externally imposed hydrodynamical or mechanical forces cause the solid–liquid interfacial area to increase beyond static equilibrium. Forced wetting plays an essential role in industrial coating processes where a thin layer of liquid is deposited continuously onto a moving solid surface. Failure of the liquid to displace sufficient air from the substrate limits the speed at which a coating can be applied. Forced wetting is also important in polymer processing and enhanced oil recovery.

Spontaneous spreading, on the other hand, is the advancing of a liquid on a solid surface towards thermodynamic equilibrium. The driving force is not imposed externally, but stems from liquid–solid interactions. Spontaneous spreading is of practical relevance in the application of paints, adhesives, lubricants, in detergency, flotation, for the migration of inks, etc. Dynamical aspects of wetting are far from being understood, which is certainly due to their inherent complexity. Good introductions to the subject are Refs. [275, 276].

Several experimental arrangements are used to measure and analyze the wetting of liquids on solid surfaces. Typical geometries are a spreading drop on a solid surface, liquid–fluid displacement through a capillary tube, steady immersion or withdrawal of fibers, plates or tapes from a pool of liquid, and the rotation of a horizontal cylinder in a liquid (Fig. 7.12).

Apparent contact angles depend on the speed of wetting. When the three-phase contact line is made to advance with a certain speed, the apparent contact angle $\Theta_{app}$ becomes higher than the advancing contact angle just before the wetting line starts to advance $\Theta_{adv}$ (Fig. 7.13). Similarly, when the wetting line is made to recede, then $\Theta_{app}$ becomes less than $\Theta_{rec}$.

Experiments on forced wetting showed that, in general, the apparent contact angle depends not only on the speed $v$ but also on the viscosity $\eta$, and the surface tension $\gamma_L$ of the liquid.

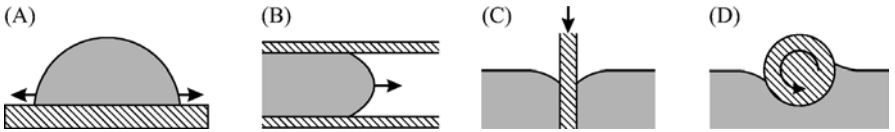



**Figure 7.12:** Some geometries used to study the dynamics of contact angle phenomena: (A) spreading drop, (B) liquid advancing in a capillary, (C) plate, tape or fibre immersed in a liquid, and (D) rotating cylinder in a pool of liquid.

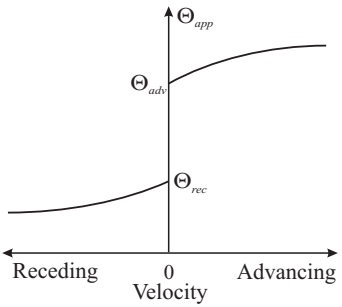



**Figure 7.13:** Schematic velocity dependence of the experimentally determined apparent contact angle $\Theta_{app}$.

Often these are correlated and the contact angle is affected only by one parameter, the so called **capillary number**:

$$Ca \equiv \frac{v\eta}{\gamma_L} \tag{7.32}$$

Especially for low capillary numbers the apparent contact angle is a function of $Ca$ only [277]. For this reason, results on dynamic wetting are usually plotted with respect to the capillary number.

■ **Example 7.6.** To study the contact angle at different velocities Blake et al. [278] drew a film of poly(ethylene terephthalate) (PET) vertically into a glass tank filled with a water/glycerol mixture. This is the commonly used plunging-tape experiment. Dynamic contact angles $\Theta_{app}$ were determined either directly, with an optical goniometer, or from high-speed video images. Figure 7.14 shows results obtained for two mixtures: 16% glycerol ($\eta = 0.0015$ Pa s, $\gamma_L = 0.0697$ N/m, $\Theta_{adv} = 72.5°$) and 59% glycerol ($\eta = 0.010$ Pa s, $\gamma_L = 0.0653$ N/m, $\Theta_{adv} = 64.5°$). At high wetting speeds both graphs increase monotonically up to a contact angle of 180°. At such high capillary numbers a layer of air is entrained between the liquid and the solid. This intermediate layer of air is unstable and at some point the liquid eventually gets into direct contact with the solid. Despite the fact that the viscosities differ by a factor of 7 the curves agree reasonably well up to capillary numbers of 0.1. The difference at low capillary numbers is due to the fact that the advancing contact angles differ by 8°.

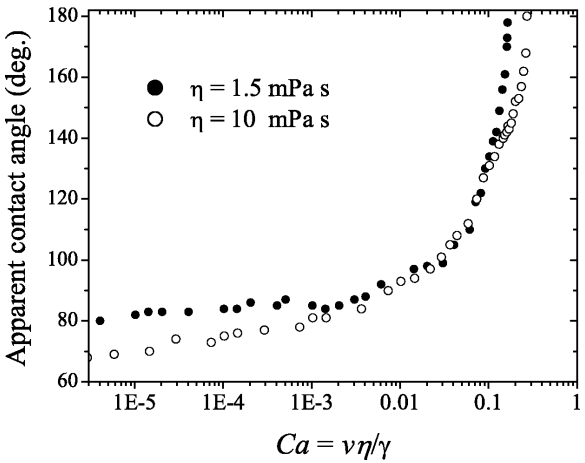



**Figure 7.14:** Apparent advancing contact angles versus the capillary number for two glycerol/water mixtures on a film of PET. Redrawn after Ref. [278].

When analyzing wetting it is important to know whether we are in a saturated vapor or if we are working under "dry" conditions. In the first case the solid surface is usually covered with a molecularly thin film of adsorbed liquid molecules and the drop spreads on this adsorbed film. In the "dry" case the solid surface is usually not covered by molecules of the liquid and is not in equilibrium.

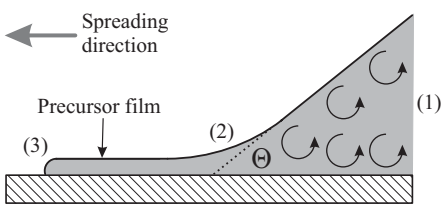



**Figure 7.15:** Liquid spreading on a solid surface with a precursor film .

The schematic structure of an advancing drop is shown in Fig. 7.15. We can classify the wetting processes according to the dissipated energy and we can distinguish three regimes: (1) The movement in the thick, wedge-shaped part of the drop, (2) the precursor film, and (3) the binding of molecules directly at the front of the precursor film . The movement in the thick, wedge-shaped part of the drop can be understood based on classical hydrodynamics. Small eddies cause a net transport of liquid in the spreading direction. Heat is dissipated due to viscous friction.

For systems which spread spontaneously it is well-known that a spreading drop forms a thin ($< 0.1\,\mu$m) primary or precursor film [279–282]. Its thickness and extension are determined by surface forces. In the precursor film, energy is dissipated by viscous friction. The liquid transport in the precursor film is driven by the disjoining pressure in the precursor film which sucks liquid from the wedge of the drop.

■ **Example 7.7.** Cazabat et al. observed the spreading of polydimethylsiloxane (PDMS) on a cleaned silicon wafer [283]. PDMS is a nonvolatile polymer which is liquid at room temperature. The PDMS used had a mean molecular weight of 9.7 kg/mol and a viscosity of 0.2 Pa s. Spreading was observed by an ellipsometer which had a lateral resolution of of 30 $\mu$m. In Fig. 7.16 we can clearly see that, in front of the drop, there is a precursor film of $\approx$ 0.7 nm thickness.

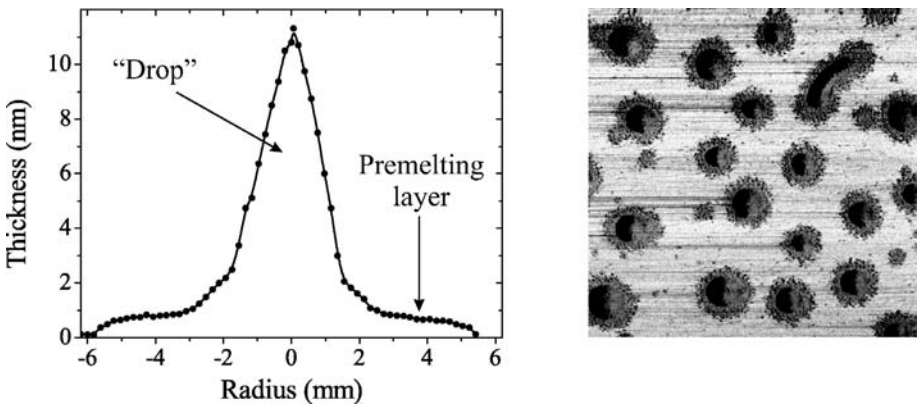



**Figure 7.16:** Left: Spreading of a drop of PDMS on a silicon wafer observed with an ellipsometer 19 h after deposition. Redrawn after Ref. [283]. On a much smaller scale (right) prewetting layers around droplets of polystyrene on a flat silicon oxide surface are observed. The atomic force microscope image shows an area of (2.5 $\mu$m)$^2$ [284].

■ **Example 7.8.** Steinhart et al. [288] used the formation of a precursor film to fabricate nanotubes of polymers. For this purpose a porous alumina filter with cylindrical holes of defined size is brought into contact with a polymer melt (Fig. 7.17). This is done at a temperature, where the polymer is liquid, which is significantly above room temperature. The precursor film of the polymer wets the walls of the pores within, typically a few seconds. Before the pores are filled completely the filter is removed from the melt, the sample is cooled to room temperature, and the filter material is dissolved in potassium hydroxide.

How does wetting of a dry solid surface at the end of the real contact line take place? In front of the precursor film molecules adsorb to the solid surface. During this adsorption heat is released. This process is probably determined by statistical kinetics on the molecular level in which the adsorbing molecules have to find free binding sites on the surface. Several microscopic models have been proposed which are based on computer simulations [285–287].

One theory, the molecular-kinetic theory suggested by Blake [289], illustrates the method of analysis (Fig. 7.18). It has the advantage that it is still analytically solvable. In this theory it is assumed that liquid molecules bind to adsorption sites on the solid surface. They must in part displace adsorbed gas molecules. These adsorption and desorption processes take place statistically, with a higher probability in the advancing direction. Thus, the spreading velocity

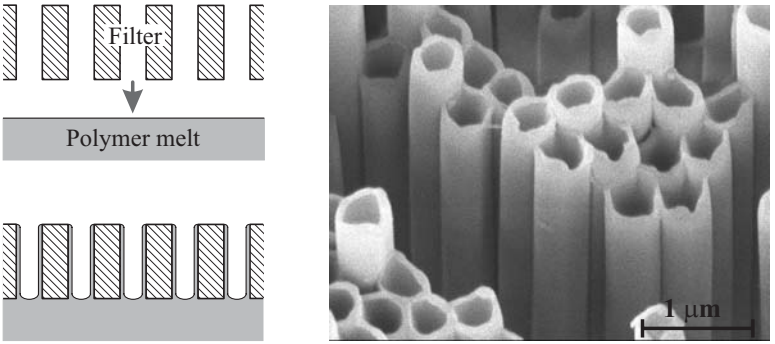



**Figure 7.17:** Nanotubes fabricated by letting a precursor film of polystyrol enter the cylindrical pores of filter [288]. A schematic of the two main steps in the fabrication process (left) and a scanning electron microscope image (right) are shown. Thanks to M. Steinhard for providing us with the picture.

of a liquid depends on the disturbance of the adsorption equilibria. At the rear of the contact line the molecules can still rearrange and reorganize.

The molecular-kinetic theory predicts a maximum wetting speed $v_{max}$ and a minimum dewetting speed $v_{min}$. At speeds larger than $v_{max}$ gas bubbles form. This was indeed observed. For water this value is $v_{max} \approx 5 - 10$ m/s.

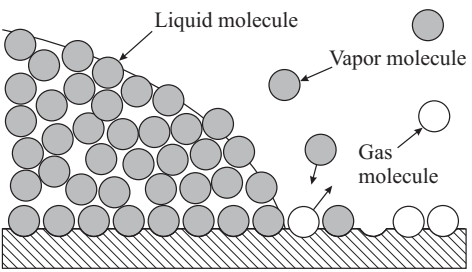



**Figure 7.18:** Illustration of the molecular-kinetic model.

De Gennes and Cazabat proposed an alternative continuum model [290]. They describe the spontaneous spreading of a liquid by a competition between the driving force, which is the disjoining pressure in the precursor film and the core region, and the friction between layers of liquid with the solid [287].

### 7.5.2 Dewetting

In many applications, films on solid surfaces are only metastable or they are stable only from a certain thickness. Example are metal films prepared by evaporation [291] and many polymer films. Since most paints and coatings contain a substantial amount of polymer, this is a highly relevant case.

How can we produce polymer films? First, the polymer is dissolved in a suitable, volatile solvent. Then two methods are widely used:

- **Dip coating** [292]. The plate to be coated is immersed into a pool containing the solution. Then it is removed at a certain withdrawal speed $v$. Excess solution is allowed to drain into the pool. The remaining solvent evaporates, leaving the polymer film behind on the plate. To remove the remaining solvent the film has to be annealed or kept overnight in a vacuum chamber. Thicknesses between 20 nm and 50 $\mu$m can be made by dip coating. Dip coating is widely applied for making anti-reflective coatings or to paint objects.

- **Spin coating** [293, 294]. The solid in the form of a flat disc is placed horizontally on a rotor. The solution is dropped onto the solid and the rotor is switched on. Due to the centrifugal force the liquid wets the whole disk. After the solvent has evaporated, a polymer film is left on the surface. Using spin coating very thin, homogeneous films can be made and the thickness can be adjusted with good precision. Spin coating is, for instance, used to cover silicon wafers with photoresist in microcircuit fabrication.

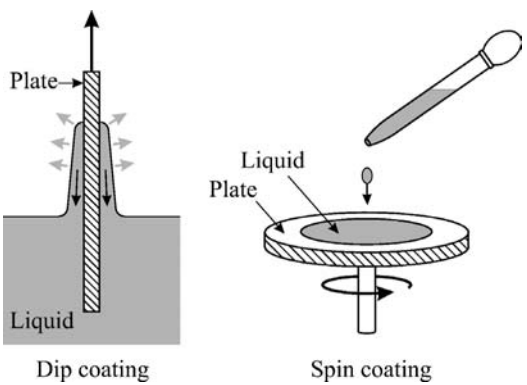



**Figure 7.19:** Dip and spin coating are two common techniques to form polymer films on solid surfaces.

When the polymer wets the solid ($\Theta = 0$), polymer films are thermodynamically stable. For $\Theta > 0$ the films are only metastable. When a thin, metastable film is heated above the glass transition temperature, holes start to form spontaneously, usually at small defects. The holes increase in size until only a network of polymer lines is formed which eventually breaks up into individual droplets (Fig. 7.19) [150]. The film stability of films with a thickness of 1–100 nm is determined by long-range surface forces, mainly van der Waals forces [151, 268, 294, 295].

## 7.6 Applications

### 7.6.1 Flotation

Flotation (review: Refs. [296, 297]) is a method to separate various kinds of solid particles from each other. It is of enormous importance to the mining industry where it is used for large-scale processing of crushed ores. The desired mineral is separated from the gangue or

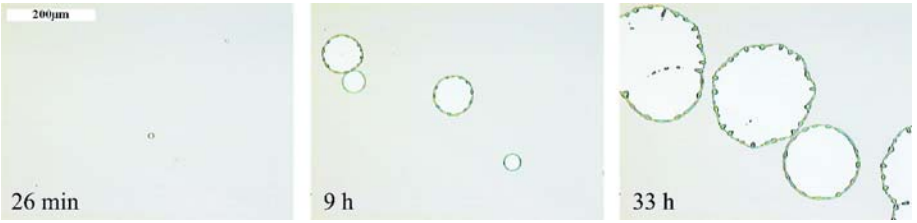



**Figure 7.20:** Dewetting of a polystyrene film of 28 nm thickness from a silicon wafer at 121°C. Thanks to C. Lorenz, P. Müller-Buschbaum, and M. Stamm.

non-mineral containing material. Originally the procedure was applied only to some sulfides and oxides (iron oxides, rutile, quartz). Meanwhile many minerals such as gold, borax, pyrite, phosphate minerals, fluorite, calcite, and apatite are separated by flotation. Another large application is the removal of unwanted material for water purification and to clean industrial waste products. The de-inking of paper is based on a very similar process .

In flotation of ore the material is first crushed to under 0.1 mm particle size. The particles are mixed with water and form a sol. This sol is called pulp. The pulp flows into a container and air bubbles are passed through (Fig. 7.21). The mineral rich particles bind to the air bubbles by hydrophobic forces and are carried to the surface of the container. A stable foam, also called froth, is formed. With the froth the mineral-rich particles can be skimmed off and removed.

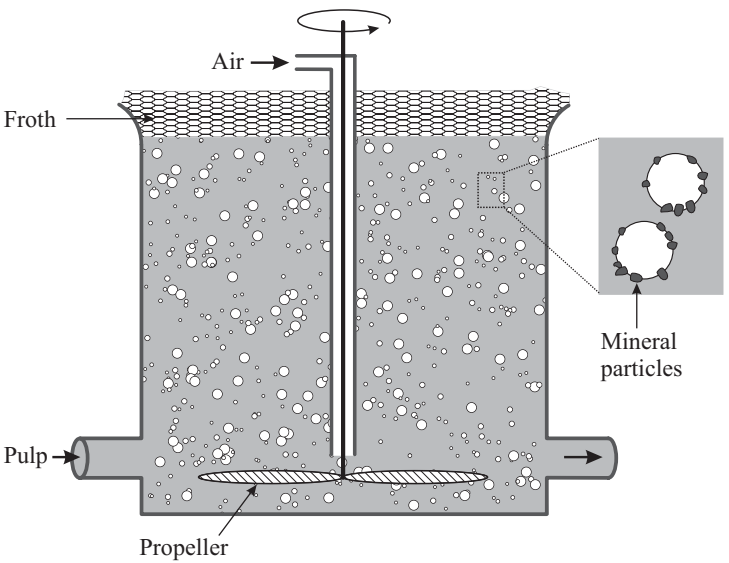



**Figure 7.21:** Schematic of a flotation system. The bubbles are usually much larger than the particles so that the water–air interface is almost planar with respect to the particles.

The wetting properties of the particles play a crucial role in flotation. We have already discussed the equilibrium position of a particle in the water–air interface (Section 7.2.2). The higher the contact angle the more stably a particle is attached to the bubble (Eq. 7.19) and the more likely it will be incorporated into the froth. Some minerals naturally have a hydrophobic surface and thus a high flotation efficiency. For other minerals surfactants are used to improve the separation. These are called collectors, which adsorb selectively on the mineral and render its surface hydrophobic. Activators support the collectors. Depressants reduce the collector's effect. Frothing agents increase the stability of the foam.

### 7.6.2 Detergency

Detergency is about the theory and practice of the removal of foreign material from solids by surface-active substances. This definition excludes pure mechanical cleaning. Also a pure chemical cleaning, e.g., by solvation of the foreign material, is not considered. In textiles oily substances usually attach to the fibres (animal fats, fatty acids, hydrocarbons, etc.). Also dust, soot, and other solid particles have to be removed in a washing process. In order to test the effectiveness of a surfactant, textiles are often polluted with standard dirt mixtures and cleaned with a standard washing procedure (launderometer). Often cleanliness is measured on the basis of optical reflectivity of white textiles.

Dirt particles spontaneously leave a solid surface if it is energetically favorable to replace the dirt–solid interface (SD) by two interfaces: the dirt–aqueous solution interface (DW) and the solid–aqueous solution interface (SW). Here, the solid is a textile fibre or any other material, which we want to clean (Fig. 7.22). The change in the Gibbs free energy

$$\Delta G = A \cdot (\gamma_{DW} + \gamma_{SW} - \gamma_{SD}) \leq 0 \tag{7.33}$$

must be negative. $A$ is the contact area. This condition can be simplified to

$$\gamma_{SD} \geq \gamma_{DW} + \gamma_{SW} \tag{7.34}$$

In this estimation we have assumed that the shape of the dirt does not change upon removal.

For liquid dirt, like grease or oil, the contact angle must become zero, i.e. the spreading coefficient should be zero or positive. This leads to the same condition and we can formulate a general requirement for a good surfactant: An effective surfactant should reduce $\gamma_{SW}$ and $\gamma_{DW}$, without reducing $\gamma_{SD}$ too strongly. A decrease in the surface tension of water — visible by bubble formation — is not a proof for an effective surfactant for detergency.

An important characteristic of a surfactant is its ability to keep dirt particles in solution (suspending power). Without this ability a washing procedure would only lead to a uniform distribution of the dirt. Obviously, the surfactant is bound to the surface of the dirt particles and keeps them dispersed. Aggregation and flocculation is prevented, for instance, by electrostatic repulsion. Referring to liquid dirt it was originally assumed that the cleaning effect of surfactants is caused by their ability to accommodate hydrophobic substances inside micelles (surfactant aggregates, see Section 12.2). However, this does not seem to be the dominating factor because the cleaning ability already increases with surfactant concentration at concentrations below the CMC , thus before micelles are formed.

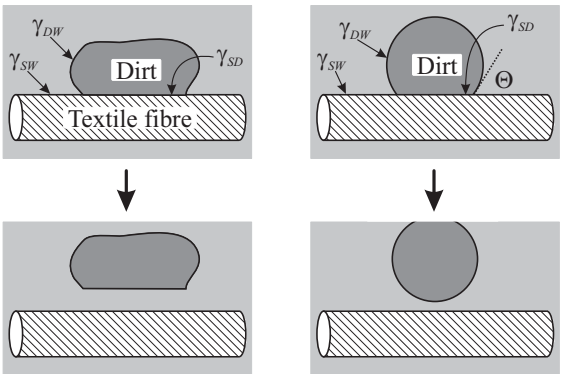



**Figure 7.22:** Removal of solid (left) and liquid dirt (right) from a textile fibre.

### 7.6.3 Microfluidics

Microfluidics is about the flow of tiny amounts of liquids. The prefix "micro" indicates that at least in two dimensions, the liquid should be confined in micrometer dimensions. If we are, for instance, dealing with a channel, its width or diameter should be below 100 $\mu$m to earn the title "microchannel". Please note that a microliter is a relatively large volume in microfluidics since it is equal to the volume of (1 mm)$^3$.

Microfluidics has become an important field of research and development mainly driven by biological and analytical applications [298]. It allows us to analyze small amounts of substances in a highly parallel way. For this reason it is widely used in DNA and protein analysis [299, 300]. Miniaturization is also attractive in chemical synthesis and significant effort is put into making chemical reactions on a microchip (lab-on-chip). Advantages are that high concentrations can be reached even with tiny amounts of substances. Heat is dissipated fast so that reactions occur practically at constant temperature. This reduces the danger of overheating and explosions. The flow is always laminar because the Reynolds number

$$Re = \frac{\rho v d}{\eta} \tag{7.35}$$

is much lower than one. Here, $\rho$ is the density of the liquid, $\eta$ is its viscosity, $v$ is its average velocity, $d$ is a typical dimension of the fluid confinement.

When trying to manipulate fluids on the micron scale, formidable problems have to be overcome [301]. Miniature pumps, valves, switches, and new analytical tools have to be developed. To illustrate this we discuss the fundamental problem of transporting a liquid through a capillary tube. In the macroscopic world we would apply a pressure $\Delta P$ between the two ends. According to the law of Hagen–Poiseuille the volume of liquid $V$ transported per time $t$ is (assuming laminar flow)

$$\frac{V}{t} = \frac{r_C^4 \Delta P}{8\eta L} \tag{7.36}$$

Here, $r_C$ is the radius and $L$ the length of the capillary tube. We see that, due to the strong dependence on the radius, we would have to apply enormous pressure to keep a significant

flow. For this reason usually electro-osmosis is used to drive a flow of aqueous solutions in microfluidics.

■ **Example 7.9.** A simple microfluidic electrophoretic device is shown in Fig. 7.23. The device allows us to separate different platinum complexes used in anti-cancer therapy [302]. In our case, five platinum complexes were separated. The electrophoretic device consists of reservoirs for the sample (A), buffer (B), and waste (C) and (D). The channels in quartz capillaries were 70 $\mu$m in diameter. The aqueous solution contained 50 mM SDS, 25 mM $Na_2B_4O_7$, 50 mM $NaH_2PO_4$ at pH 7.0 in addition to the analyte. Applying a positive voltage at (A) and a negative voltage at (C) leads to an electro-osmotic flow from (A) to (C) because quartz ($SiO_2$) is negatively charged at neutral pH (see Table 5.1). A small plug of sample solution is drawn into the separation channel (BD). Then the voltage between (A) and (C) is switched off and a voltage between (B) and (D) is applied. (B) is positively charged, (D) is negative. The platinum complexes are carried with the osmotic flow towards (D). Superimposed is the electrophoretic drift which separates the platinum complexes. Unfortunately platinum complexes are usually not charged in an aqueous medium and no electrophoresis occurs. For this reason the author used a trick and added SDS. The platinum complexes are enclosed in negatively charged SDS micelles. Depending on their precise structure the micelles have a different size or charge. In this way a mixture can be separated into its components. Molecules were detected by optical adsorption at 210 nm.

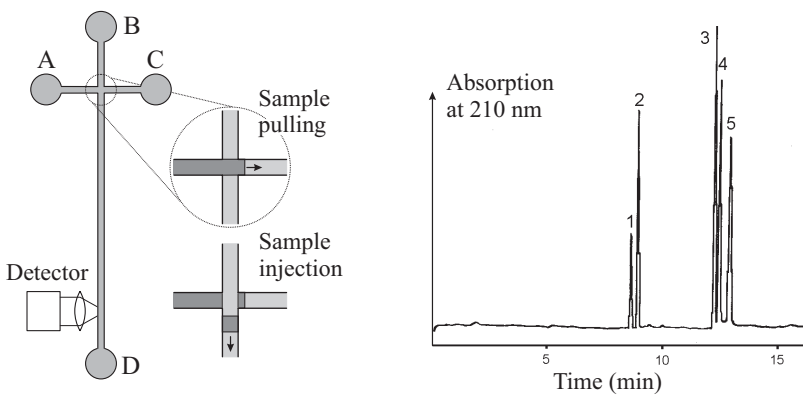



**Figure 7.23:** Simple electrophoretic microfluidic system (see example 7.9).

## 7.6.4 Adjustable wetting

For many applications it is desirable to be able to adjust the wetting properties of a solid surface for aqueous solutions. One method is called electrowetting [303, 304]. In electrowetting an electric potential is applied between a metal surface and a liquid via an electrode (Fig. 7.24). The metal is coated with an insulating layer of thickness $h$. Fluoropolymer coatings turned out to be suitable materials though it is not fully understood why [305]. The change in contact

angle is described by

$$\cos\Theta = \cos\Theta_0 + \frac{\varepsilon\varepsilon_0 U^2}{2h\gamma_L} \tag{7.37}$$

Here, $\Theta_0$ is the contact angle without applied voltage, $\varepsilon$ is the dielectric permittivity of the insulating polymer film, and U is the applied voltage. Please note that the sign of the applied potential is not relevant.

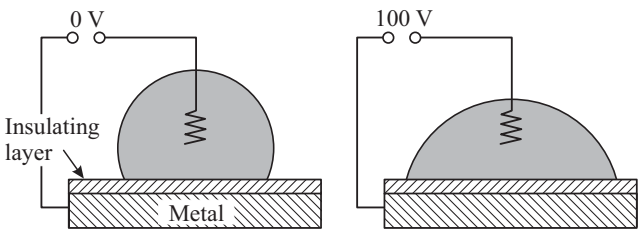



**Figure 7.24:** Electrowetting of a liquid drop on a metal which is coated with a thin insulating layer.

■ **Example 7.10.** Prins et al. [306] used electrowetting to control fluid motion in microchannels. To do so, they coated aluminum electrodes first with a 12 $\mu$m thick layer of parylene and then with a 10 nm thick fluoropolymer film. The channels were 0.35 mm wide. Due to the hydrophobic polymer water does not flow into the capillaries. Only after applying voltages of typically 200 V did the capillaries fill with water. When switching the voltage off, the water flowed out of the capillaries again.

Another approach to control wetting is to bind molecules covalently to a surface which can undergo reversible conformational transitions between a hydrophobic and a hydrophilic state. To switch between conformations, light azobenzene is an attractive candidate to use. Azobenzene assumes either a cis- or trans-configuration. Illumination by UV light (350 nm wavelength) leads to a trans-cis photoisomerisation and the concentration of cis-azobenzene is increased. Blue light (450 nm) initiates the reverse reaction and leads to an accumulation of trans-azobenzene. By attaching appropriate rest groups the wetting behavior of an azobenzene layer can be switched by shining either blue or UV light onto the surface [307].

■ **Example 7.11.** Langer et al. [308] formed monolayers of (16-mercapto) hexadecanoic acid ($HS(CH_2)_{15}COOH$) on gold. With their thiol group ($\sim$SH) these molecules bind spontaneously and strongly to the gold surfaces. To allow for conformational changes the molecules are not densely packed but a certain distance is kept between them. In water, the hydrophilic carboxy groups dissociate ($\sim COOH \rightleftharpoons \sim COO^- + H^+$). When applying a negative or neutral potential to the gold the molecules stand upright and expose their hydrophilic group to the water. This leads to a low contact angle. At positive potentials the carboxyl groups bend towards the gold surface thus exposing the hydrophobic alkyl chains to the aqueous medium. This leads to a high contact angle. By applying the appropriate potential the contact angle could be adjusted between 15° and 45°.

## 7.7 Summary

- Young's equation relates the contact angle to the interfacial tensions $\gamma_S$, $\gamma_L$, and $\gamma_{SL}$:

$$\gamma_L \cdot \cos\Theta = \gamma_S - \gamma_{SL}$$

- The work required to create a new three-phase contact line per unit length is called line tension. It is typically of the order of 0.1 nN. For tiny liquid drops the line tension can significantly influence the wetting behavior.
- A liquid rises in a lyophilic ($\Theta < 90°$) capillary. The height increases with decreasing capillary radius. From lyophobic ($\Theta > 90°$) capillaries a liquid is expelled.
- Contact angles are commonly measured by the sessile drop, the captive bubble, and the Wilhelmy plate method. To characterize the wetting properties of powders the capillary rise method is used.
- Contact angle hysteresis can be caused by surface roughness, heterogeneity, dissolved substances, and structural changes of the solid at the three-phase contact line.
- For liquid thicknesses up to 10–100 nm, surface forces influence the wetting behavior directly. This is, for instance, the case in the core region of the three-phasecontact line and for thin films.

## 7.8 Exercises

1. In a glass capillary of 1 mm diameter water rises 2.2 cm at 25°C. What is the difference in the surface tension between the glass–vapor and the glass–liquid interface?
2. A small drop is placed on a solid surface. Its contact angle is 90° for a contact radius of 0.5 mm. Its line tension is $\kappa$ = 0.5 nN, its surface tension 50 mN/m. How does the contact angle change depending on the contact radius $a$? How does the contact angle change for a negative line tension of $-0.5$ nN? Please plot $\Theta(a)$ for both cases.
3. In a classical paper Owens & Wendt estimated the surface energy of polymers by measuring the contact angle with water and methylene iodide [272]. They measured:

| | $\Theta$ with water | $\Theta$ with methylene iodide |
|---|---|---|
| Polyethylene | 94° | 52° |
| Poly (vinyl chloride) | 87° | 36° |
| Polystyrene | 91° | 35° |
| Poly (methyl methacrylate) | 80° | 41° |

The dispersive and polar components of the surface tensions of the liquids were estimated to be $\gamma_L^d = 21.8$ mN/m and $\gamma_L^p = 51.0$ mN/m for water and $\gamma_L^d = 49.5$ mN/m and $\gamma_L^p = 1.3$ mN/m for methylene iodide. This estimation was done by measuring contact angles with various hydrocarbons and assuming that there are only nonpolar interactions. What are the surface energies, $\gamma_S$, of the polymers?

# 8 Solid surfaces

## 8.1 Introduction

Molecules at a surface are arranged in a different way from molecules in the bulk. This statement is, in general, true for solids and liquids. There is, however, a major difference: when a solid is deformed by small external forces it reacts elastically while for a liquid the reaction is plastic. The shape of the solid was generated in the past and it is usually not determined by the surface tension. Individual atoms and molecules are only able to vibrate around their mean position and are otherwise fixed to a certain site.

This applies, however, only within limits. Many solids are somewhat mobile and can flow very slowly. In that case methods and models of capillarity can be applied. One case where capillarity plays an important role is sintering. In sintering a powder is heated. At a temperature of roughly 2/3 of the melting point of the material the surface molecules become mobile and can diffuse laterally. Thereby the contact areas of neighbouring particles melt and menisci are formed. When cooling, the material solidifies in this new shape and forms a continuous solid.

In this chapter we mainly deal with the microscopic structure of crystalline solid surfaces and with the methods used to analyze this structure and the chemical composition (introductions are Refs. [309–311]).

Most solids are not crystalline on their surface. This is certainly true for amorphous solids. It is also true for most crystalline or polycrystalline solids because for many materials the molecular structure at the surface is different from the bulk structure. Many surfaces are for example oxidized under ambient conditions. A prominent example is aluminum which forms a hard oxide layer as soon as it is exposed to air. Even in an inert atmosphere or in ultrahigh vacuum (UHV) the surface molecules might form an amorphous layer on the crystalline bulk solid.

Since most natural surfaces are amorphous, it may seem somewhat academic to study crystalline surfaces. But there are good reasons to do so:

- The well defined structure of crystalline surfaces allows comparison of experiments on different samples of the same material.
- The periodic structure of crystalline surfaces facilitates theoretical description and allows us to use powerful diffraction methods to analyze them.
- Crystalline surfaces are important in the semiconductor industry. Many modern semiconductor devices depend on the defined production of crystalline surfaces.

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

For the investigation of clean crystalline surfaces ultrahigh vacuum has to be used. To keep a surface clean for time periods of hours, gas pressures lower than $10^{-7}$ Pa are required (see Exercise 2.1). To provide such conditions is quite demanding. A textbook on vacuum technology is Ref. [312] and a review on history and current developments of vacuum technology is Ref. [313].

## 8.2 Description of crystalline surfaces

### 8.2.1 The substrate structure

Let us start with the simple case of an ideal crystal with one atom per unit cell that is cut along a plane, and assume that the surface does not change. The resulting surface structure can then be described by specifying the bulk crystal structure and the relative orientation of the cutting plane. This ideal surface structure is called the *substrate structure*. The orientation of the cutting plane and thus of the surface is commonly notated by use of the so-called Miller indices.

Miller indices are determined in the following way[1] (Fig. 8.1): The intersections of the cutting plane with the three crystal axes are expressed in units of the lattice constants. Then the inverse values of these three numbers are taken. This usually leads to non-integer numbers. All numbers are multiplied by the same multiplicator to obtain the smallest possible triple of integer numbers. The triple of these three numbers $h$, $k$, and $l$ is written as $(hkl)$ to indicate the orientation of this plane and all parallel planes. Negative numbers are written as $\bar{n}$ instead of $-n$. The notation $\{hkl\}$ is used specify the $(hkl)$ planes and all symmetrical equivalent planes. In a cubic crystal, for example, the (100), (010), and (001) are all equivalent and summarized as $\{100\}$.

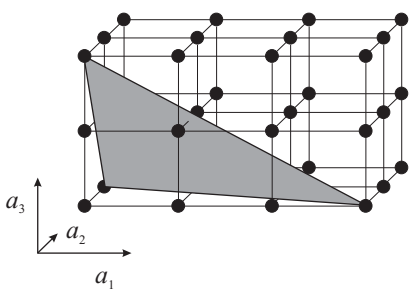



**Figure 8.1:** Notation of a cutting plane by Miller indices. The three-dimensional crystal is described by the three-dimensional unit cell vectors $\vec{a_1}$, $\vec{a_2}$, and $\vec{a_3}$. The indicated plane intersects the crystal axes at the coordinates $(3, 1, 2)$. The inverse is $(\frac{1}{3}, \frac{1}{1}, \frac{1}{2})$. The smallest possible multiplicator to obtain integers is 6. This leads to the Miller indices $(263)$.

[1] A more general definition of the Miller indices is given in the appendix on diffraction in the context of the reciprocal lattice.

Hexagonal close packed crystals are often characterized using four lattice vectors. In this case four Miller indices ($h\ k\ i\ l$) are used correspondingly, where the fourth index is related to the first two indices by $i = -(h + k)$.

In surface science, often low-index surfaces, i.e. crystal surfaces with low Miller indices, are of special interest. In Fig. 8.2 the three most important low-index surfaces of a face centered cubic lattice are shown. The (100) is equivalent to the (010) and the (001) surfaces, the (110) is equivalent to the (011) and the (101) surfaces.

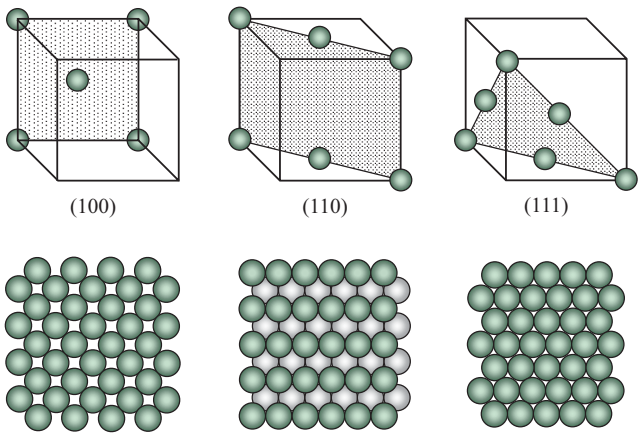



**Figure 8.2:** Low index surfaces for a face centered cubic crystal.

Crystalline surfaces can be divided into five Bravais lattices (Fig. 8.3) according to their symmetry. They are characterized by the lattice angle $\alpha$ and the lengths of the lattice vectors $a_1$ and $a_2$. The position vectors of all individual surface atoms can be indicated by

$$\vec{r} = n \cdot \vec{a}_1 + m \cdot \vec{a}_2 \tag{8.1}$$

where, $n$ and $m$ are integers.

### 8.2.2 Surface relaxation and reconstruction

Atoms at solid surfaces have missing neighbors on one side. Driven by this asymmetry the topmost atoms often assume a structure different from the bulk. They might form dimers or more complex structures to saturate dangling bonds. In the case of a **surface relaxation** the lateral or in-plane spacing of the surface atoms remains unchanged but the distance between the topmost atomic layers is altered. In metals for example, we often find a reduced distance for the first layer (Table 8.1). The reason is the presence of a dipole layer at the metal surface that results from the distortion of the electron wavefunctions at the surface.

If the lateral distances of the atoms are changed, this is called **surface reconstruction**. This is, for instance, observed with the (100) faces of Au, Ir, Pt, and W. Figure 8.4 shows two types of surface reconstruction that lead to a doubling of the lattice spacing in one direction. Semiconductor surfaces tend to exhibit surface reconstruction due to the directional character of the dangling bonds at the surface.

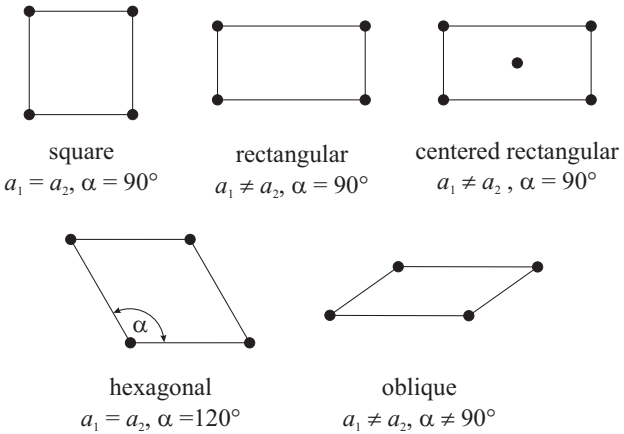



**Figure 8.3:** The five two-dimensional Bravais lattices.

**Table 8.1:** Reduction of the interlayer spacing between the two top layers as compared to the bulk due to surface relaxation for the (110) surfaces of unreconstructed Cu, Ni, Au, and Pd.

| Cu | Ni | Au | Pb |
|---|---|---|---|
| 0.020 Å | 0.156 Å | 0.125 Å | 0.080 Å |

For reconstructed, crystalline surfaces the position vectors of the atoms have to be described by new unit cell vectors $\vec{b_1}$ and $\vec{b_2}$:

$$\vec{r} = n' \cdot \vec{b_1} + m' \cdot \vec{b_2} \tag{8.2}$$

Often there is a relationship between $\vec{a_1}, \vec{a_2}$ and $\vec{b_1}, \vec{b_2}$ of the form

$$\vec{b}_1 = p \cdot \vec{a}_1 \qquad \text{and} \qquad \vec{b}_2 = q \cdot \vec{a}_2 \tag{8.3}$$

with $p$ and $q$ being integers. The surface structure is denoted in the form

$$A(hkl)(p \times q)$$

where $A$ is the chemical symbol of the substrate.

■ **Example 8.1.** Gold (111) is an example of a surface without reconstruction. It is denoted as Au(111)$(1 \times 1)$. The (100) surface of silicon exhibits a $(2 \times 1)$ reconstruction and is denoted as Si(100)$(2 \times 1)$. The $7 \times 7$ surface reconstruction of Si(111) (Fig. 8.5) is another example for a complex surface reconstruction [329, 330].

In general, the surface structure can be more complex as every lattice site may be occupied by more than one atom. This can be described by the so-called basis that indicates the fixed relative orientation of the atoms or molecules. In this case the lattice type and the basis have to be given for a complete description of the crystal structure. In addition, it is possible that the surface atoms lie at different heights (Fig. 8.6). Then the basis has to be extended to more than just the surface layer of atoms.

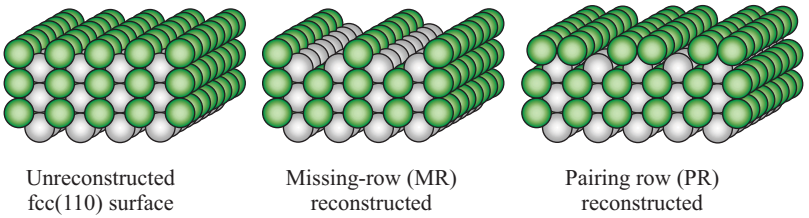



**Figure 8.4:** Typical reconstructions of face centered cubic (110) surfaces.

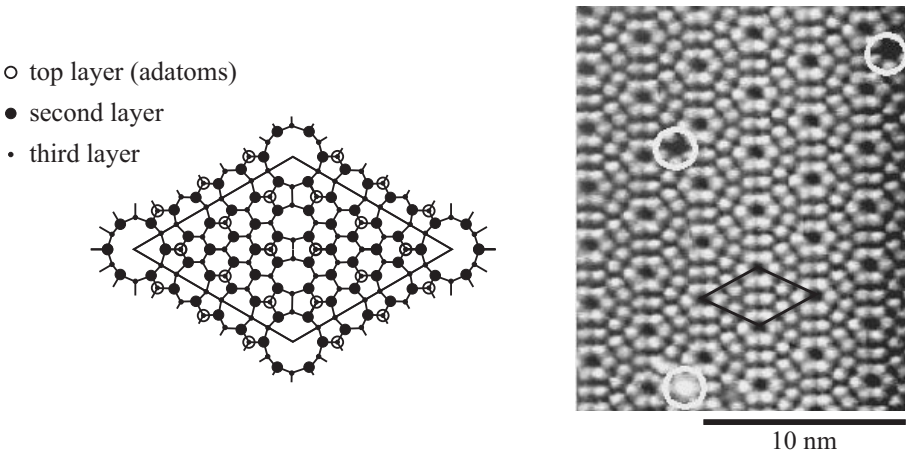



**Figure 8.5:** a) Schematic of the 7×7 reconstruction of the Si(111) surface. b) Image of a Si 7×7 surface obtained with a combination of STM and a special non-contact AFM mode [331] (see Section 8.6.3). Only the adatoms are visible under the chosen imaging conditions. Single defects of the surface structure are resolved (circles). The picture was kindly provided by E. Meyer.

### 8.2.3 Description of adsorbate structures

If molecules adsorb to a crystalline surface they often form a crystalline overlayer which is called a superlattice. This occurs if the adsorbates bind preferentially to specific sites. Then the positions of the adsorbed molecules can be written as

$$\vec{r}_{ad} = n'' \cdot \vec{c}_1 + m'' \cdot \vec{c}_2 \tag{8.4}$$

Again, $n''$ and $m''$ are integers. Usually there are fewer adsorbed molecules than there are atoms in the underlying crystal surface. Therefore the unit cell formed by the vectors $\vec{c}_1$ and $\vec{c}_2$ is often larger than the unit cell of the underlying crystal lattice described by $\vec{b}_1$ and $\vec{b}_2$. The structure of the adsorbate lattice is expressed by the ratios of the lengths of the unit cell vectors: $p' = c_1/b_1$ and $q' = c_2/b_2$. If the adsorbate lattice is rotated with respect to the underlying substrate lattice by an angle $\beta$, the value of $\beta$ preceded by an "R" is indicated.

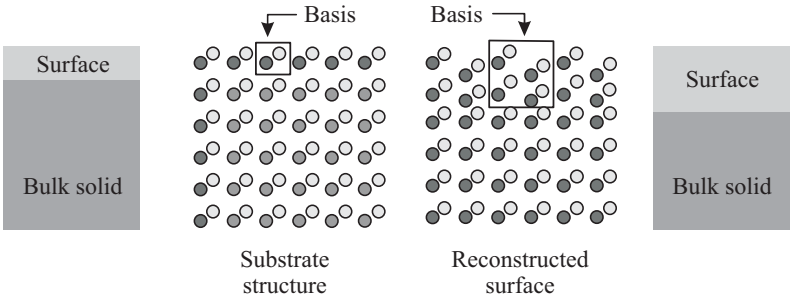



**Figure 8.6:** Example of the necessary extension of the basis in the case of surface atoms positioned at different heights.

This together leads to the so-called Wood's notation

$$A(hkl)\mathrm{c}(p' \times q')\,\mathrm{R}\,\beta - B,$$

where A and B are the chemical symbols of substrate and adsorbate, respectively. The letter "c" is added in case of a centered unit cell of the adsorbate. As examples the $(1 \times 1)$-2H structure of hydrogen on Rh(110), the $(\sqrt{3} \times \sqrt{3})$R 30° structure of chemisorbed alkanethiols on Au(111) (like Xe on graphite) and the c$(2\sqrt{2} \times \sqrt{2})$R 45° adsorption structure of CO on Pd(100) (the latter occuring at a surface coverage of $\theta = 0.5$) are shown in Fig. 8.7 [320–323].

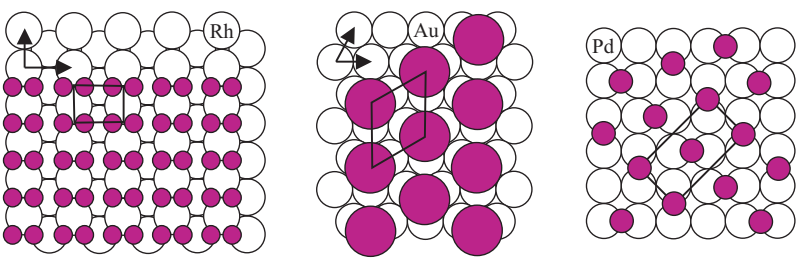



**Figure 8.7:** Adsorbate superlattices. From left to right: $\mathrm{Rh}(110)(1 \times 1) - 2\mathrm{H}$, $\mathrm{Au}(111)(\sqrt{3} \times \sqrt{3})\,\mathrm{R}\,30^\circ - \mathrm{CH_3(CH_2)}_n\mathrm{SH}$, and $\mathrm{Pd}(100)\,c\,(2\sqrt{2} \times \sqrt{2})\,\mathrm{R}\,45^\circ - \mathrm{CO}$.

## 8.3 Preparation of clean surfaces

To prepare crystalline surfaces, usually the starting material is a suitable, pure, three-dimensional single crystal. From this crystal, a slice of the desired orientation is cut. Therefore the crystal must be oriented. Orientation is measured by X-ray diffraction. Hard materials are then grounded and polished. Soft materials are cleaned chemically or electrochemically. The surfaces are still mechanically stressed, contaminated, or chemically changed, e.g. oxidized. In principle, electrochemical processes in liquid can be used to generate clean crystalline surfaces. The problem is that an electrochemical setup is not compatible with an UHV

environment. There are, however, combined ultrahigh vacuum/electrochemistry instruments (UHV-EC), where the transfer of electrochemically treated samples into the UHV chamber without contact to air, is possible (see Ref. [314]). Still the sample surfaces may undergo structural or compositional changes during the transfer. Usually *in situ* methods of surface preparation are preferred that can be applied within the UHV chamber. These methods are described in the following.

**Thermal treatment**. Heating of the material may cause desorption of weakly bound species from the surface and can therefore be used to clean surfaces. A positive side effect is that annealing reduces the number of surface defects since it increases the diffusion rates of surface and bulk atoms. There can also be some unwanted side effects: surface melting and other types of phase transitions may occur well below the bulk melting point, leading to other than the desired surface structure.

■ **Example 8.2.** To obtain a clean tungsten W(110) surface, the crystal is heated to 1600°C for about one hour in the presence of $10^{-6}$ mbar of oxygen. This oxidizes surface contaminations. A consecutive flash heating to 2000°C leads to decomposition and desorption of the oxide layer.

**Cleavage** of bulk crystals to expose clean, defined lattice planes is possible for brittle substances. Some materials, like mica or highly oriented pyrolytic graphite (HOPG) that exhibit a layered structure, are readily cleaved by just peeling off some layers or using a razor blade. For other materials the so-called "double-wedge technique" can be used (Fig. 8.8).

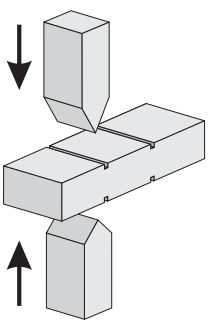


**Figure 8.8:** Double-wedge technique for cleavage of brittle materials. To ensure proper cleavage, the crystal is pre-cut and positioned between the wedges according to its crystallographic orientation.

Cleavage is a quick way to produce fresh surfaces from brittle or layered materials. There are, however, limitations: Cleavage may produce metastable surface configurations that are different from the equilibrium structure. Depending on the material, only certain cleavage planes can be realised. Crystals usually cleave along nonpolar faces, so that positive and negative charges compensate within the surface. GaAs, for example, can only be cleaved along the nonpolar {110} planes, whereas the polar {100} and {111} faces cannot be obtained.

**Sputtering**. A universal method for cleaning surfaces is sputtering. In sputtering the surface is bombarded with noble gas ions. Thereby contaminants — and in most cases the first few layers of the substrate as well — are removed. In sputtering, an inert gas (usually

argon) at a pressure of typically 1 Pa is inserted into a vacuum chamber. A high electric field is applied. In any gas and at any given moment a certain number of atoms are ionized (although very few), for example by cosmic background radiation. The electrons which are set free in the ionization process are accelerated in the electric field. They hit and ionize further gas atoms. More electrons are liberated so that eventually a plasma is created. The ions are accelerated by the applied voltage of some kV towards the sample. After sputtering, annealing is often necessary to remove adsorbed or embedded noble gas atoms and to heal defects of the crystal surface created by the bombardment.

Sputtering is an excellent and versatile cleaning technique for elemental materials. Care has to be taken when applying it to composite materials like, e.g., alloys. Sputtering rates will depend on the component and lead to changes in surface stoichiometry. In these cases, cleavage may be the better choice. The necessary annealing process can also be critical and even induce surface contamination. A prominent example is iron, that will usually contain some sulfur that segregates at the surface during annealing. With such materials, consecutive cycles of annealing and sputtering may be required to obtain a clean crystalline surface.

Sputtering cannot only be used to remove surface layers, but also to deposit material on surfaces. This process is also known as physical vapor deposition (PVD) and is often used to deposit thin metal films. For this purpose we include a target composed of the material to be deposited into the vacuum chamber (Fig. 8.9). This target again is bombarded by ions from an argon plasma. The atoms that are ejected from the target surface partially condense on our sample surface and form a thin layer. An important parameter is the kinetic energy of the $Ar^+$ ions. Low-impact energies are not sufficient to break the bonds of the target atoms and high-energy ions would penetrate deeply into the target without disrupting the surface much. Typically, energies of several hundred eV to several keV are used. An in-depth discussion of sputtering can be found in Ref. [315].

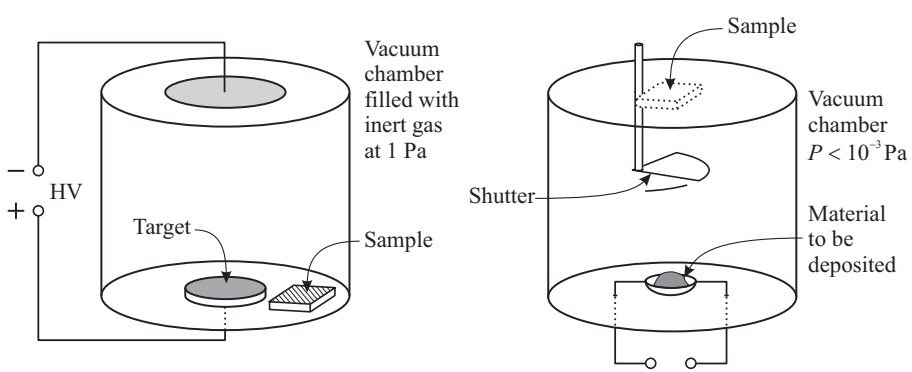



**Figure 8.9:** Schematic of a sputter (left) and an evaporation chamber (right).

**Evaporation** is a simple and well understood vacuum technique for thin film deposition. The material to be deposited is heated until it starts to evaporate. Mainly two methods are used for evaporation:

- For materials with low melting point, such as gold, the material is filled into a small heat-resistant container (evaporation boat) usually made of tungsten or tantalum. The

evaporation boat is heated up by an electrical current. Only small amounts of material can be evaporated and, if it is an alloy, the composition might change due to the different partial pressure of the single components in the alloy.

- Using an electron beam which is guided onto the material by an electromagnetic field. Using electron beam evaporation materials with high melting points which are difficult to evaporate can also be deposited.

Since evaporation is done in vacuum the mean free path of the molecules is long and they move practically in straight lines to condense onto the substrate, which is placed at an appropriate position. Typically amorphous or polycrystalline layers of some 10 nm thickness are produced by this technique. To get a crystalline surface the sample is usually annealed during or after evaporation. The film thickness is usually monitored by a quartz crystal microbalance.

In **molecular beam epitaxy** (MBE) [317], molecular beams are used to deposit epitaxial layers onto the surface of a heated crystalline substrate (typically at 500–600°C). Epitaxial means that the crystal structure of the grown layer matches the crystal structure of the substrate. This is possible only if the two materials are the same (homoepitaxy) or if the crystalline structure of the two materials is very similar (heteroepitaxy). In MBE, a high purity of the substrates and the ion beams must be ensured. Effusion cells are used as beam sources and fast shutters allow one to quickly disrupt the deposition process and create layers with very sharply defined interfaces. Molecular beam epitaxy is of high technical importance in the production of III-V semiconductor compounds for sophisticated electronic and optoelectronic devices. Overviews are Refs. [318, 319].

## 8.4 Thermodynamics of solid surfaces

### 8.4.1 Surface stress and surface tension

When describing liquid surfaces, the surface tension was of fundamental importance. If we try to extend the definition of surface tension to solids, a major problem arises [324]. If the surface of a liquid increases, then the number of surface atoms increases in proportion. For a solid surface this *plastic* increase of the surface area is not the only possible process. Usually more important is an *elastic* increase of the surface area. If the solid surface is increased by mechanically stretching, the distance between neighboring surface atoms changes, while the number of surface atoms remains constant. The change in surface area is commonly described in terms of the surface strain. The total surface strain $\varepsilon_{tot}$ is given by the change in surface area divided by the whole surface area: $d\varepsilon_{tot} = dA/A$. The surface strain may be divided into a plastic strain $d\varepsilon_p$ and an elastic strain $d\varepsilon_e$ so that $d\varepsilon_{tot} = d\varepsilon_p + d\varepsilon_e$.

To consider the elastic increase of the surface area the surface stress $\Upsilon$ is introduced. The change in the Gibbs free energy per unit area is given by the reversible work required to expand the surface against the surface tension $\gamma$ and the surface stress $\Upsilon$

$$\gamma^S = \frac{d\varepsilon_p}{d\varepsilon_{tot}} \cdot \gamma + \frac{d\varepsilon_e}{d\varepsilon_{tot}} \cdot \Upsilon \tag{8.5}$$

Here, $\gamma^S$ is called the generalized surface intensive parameter or surface energy. $\gamma^S$ is not a real thermodynamic quantity since it depends on the history of the solid [325]. It depends

on the way new surface area is created while the surface tension and the surface stress are independent of the specific process.

A relation between the surface tension $\gamma$ and the surface stress $\Upsilon$ was derived in a classical paper by Shuttleworth [326]:

$$\Upsilon = \gamma + \frac{\partial \gamma}{\partial \varepsilon_e} \tag{8.6}$$

If a surface area is increased by stretching, the increase in free energy (which in that case is $\Upsilon dA$) is given by the surface tension plus the change of the surface tension with the elastic strain. The Shuttleworth equation indicates that, in order to know the surface stress, we need to know the surface tension but also the dependence of $\gamma$ on the elastic strain $\varepsilon_e$.

There is another fundamental difference between solid and liquid surfaces. Crystals can respond differently in different directions when increasing the surface area. As a result the number of equations increases by a factor of two as we have to consider contributions for the two in-plane coordinates separately.

For many applications it is useful to know which shape a crystal would assume in equilibrium. "Equilibrium" means: if plastic deformation is allowed, for instance, during annealing. The surface tensions of a crystal are in general different for the different crystal faces. Which shape does the crystal assume for a given volume, if its entire free surface energy is minimal? A general, quasi-geometrical solution of this problem was suggested by Wulff [328]:

1. Draw a group of vectors with common starting points. The length of the vectors should be proportional to $\gamma$ of the crystal face, the direction is perpendicular (normal) to it.

2. Draw at the end of each vector, one plane perpendicular to the vector direction. The body included by these planes is the equilibrium shape of the solid.

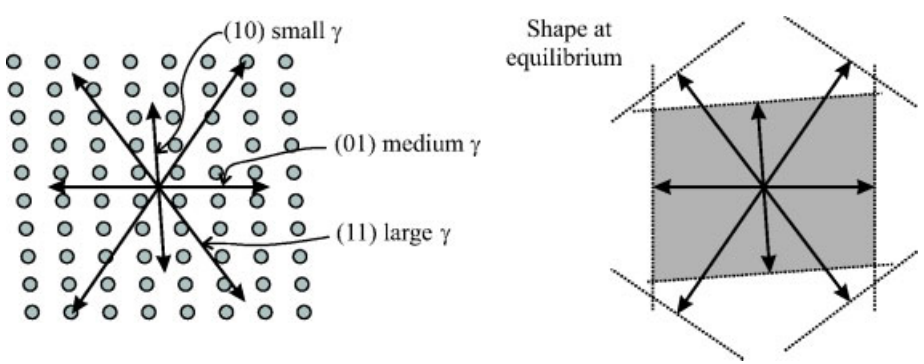



**Figure 8.10:** Wulff construction for determining the equilibrium configuration of a crystal.

### 8.4.2 Determination of the surface energy

Determining surface energy parameters such as the surface tension, the surface stress, the internal surface energy, etc., is a difficult task. This is partially due to the fact that different

methods lead to different parameters and results cannot be compared directly. In addition, solid surfaces are usually not in equilibrium. We start by discussing how surface energy parameters can be calculated.

Bonds in **covalent solids** are dominated by short-range interactions. The internal surface energy is simply calculated as half of the energy which is necessary to split the bonds that pass through a certain cross-sectional area. This is called the *nearest neighbor broken bond model*. The Gibbs free surface energy is not much different from this value because, at room temperature, the entropic contribution is usually negligible.

■ **Example 8.3.** The energy per bond between two carbon atoms in diamond is 376 kJ/mol. If we split diamond at the (111) face, then $1.83 \times 10^{19}$ bonds per $m^2$, break. Thus:

$$u^\sigma \approx \frac{376\,\mathrm{kJ\,mol^{-1}} \times 1.83 \times 10^{19}\,\mathrm{m^{-2}}}{(2 \times 6.02 \times 10^{23}\,\mathrm{mol^{-1}})} = 5.7\,\frac{\mathrm{N}}{\mathrm{m}}$$

This simple calculation demonstrates that the surface energies of solids can be much higher than the corresponding values of liquids (Table 2.1).

**Noble gas crystals** are held together by van der Waals forces. In order to calculate their surface energy we proceed as indicated in Section 6.2.3. In a thought experiment we split the crystal and calculate the required work (Fig. 8.11). This splitting at fixed atomic positions is, however, only the first step. In a second step we allow the molecular positions to rearrange according to the new situation. This is done by computer simulations. A certain interatomic potential has to be assumed. Often the Lennard–Jones potential, $V(r_{ij}) = C_1 r_{ij}^{-12} - C_2 r_{ij}^{-6}$, is used. Here, $r_{ij}$ is the distance between the atoms $i$ and $j$. The constants $C_1$ and $C_2$ depend on the material. The first term is due to the short-range repulsion of overlapping electron orbitals. It is called "Pauli repulsion". The second term represents the van der Waals dipole attraction. As a result the atoms close to the surface usually increase their distance because the attraction of second and third neighbors is missing. The surface energy is reduced.

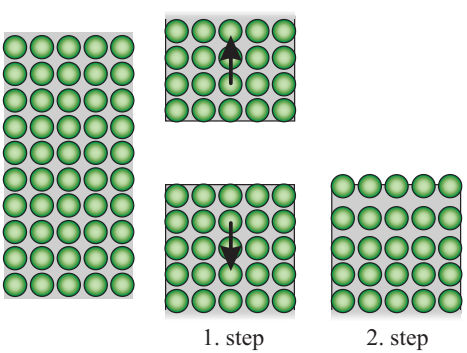



**Figure 8.11:** Splitting of a crystal: Thought experiment for the calculation of the surface energy.

Examples of the surface energies for noble gas crystals at 0 K are given in Table 8.2. The fact that a range of values is given is due to different crystal planes which leads to a variation of the surface energies $u^\sigma$.

**Table 8.2:** Internal surface energies $u^\sigma$ of noble gas crystals. A range is given because of different crystal orientations which lead to different surface energies.

| | |
|---|---|
| Ne | 19.7–21.2 mN/m |
| Ar | 43.2–46.8 mN/m |
| Kr | 52.8–57.2 mN/m |
| Xe | 62.1–67.3 mN/m |

A similar calculation can be done for **ionic crystals**. In this case the Coulomb interaction is taken into account, in addition to the van der Waals attraction and the Pauli repulsion. Although the van der Waals attraction contributes little to the three-dimensional lattice energy, its contribution to the surface energy is significant and typically 20–30%. The calculated surface energy depends sensitively on the particular choice of the inter-atomic potential.

**Table 8.3:** Calculated surface tensions $\gamma$ and surface stresses $\Upsilon$ of ionic crystals for different surface orientations compared to experimental results [325, 327]. All values are given in mN/m. (a) from cleavage experiments, (b) extrapolated from liquid.

| Crystal | orientation | $\gamma_{calc}$ | $\gamma_{exp}$ | $\Upsilon_{calc}$ | $\Upsilon_{exp}$ |
|---|---|---|---|---|---|
| LiF | (100) plane | 480 | $340^a$ | 1530 | |
| | (110) plane | 1047 | | 407 | |
| $CaF_2$ | (111) plane | | $450^a$ | | |
| NaCl | (100) plane | 212 | $300^a$,$190^b$ | 415 | 375 |
| | (110) plane | 425 | | 256 | |
| KCl | (100) plane | 170 | $110^a$,$173^b$ | 295 | 320 |
| | (110) plane | 350 | | 401 | |

For **metals** there are two methods used to calculate their surface energy. 1) As in the case for noble gases and ionic crystals the surface energy is calculated from the interaction potential between atoms. 2) Alternatively, we can use the model of free electrons in a box, whose walls are the surfaces of the metal. This quantum mechanical view is independent of the specific material. The wave functions of the electrons have nodes at the walls. When splitting the crystal some previous states of the electrons are no longer available due to the additional boundary condition. The electrons are thereby forced to occupy states of higher energy. This additional energy is the surface energy.

There are only a few and often indirect methods of measuring surface energy parameters of solids. The problems are that surfaces contaminate and that they are usually inhomogeneous. For a review see Ref. [325].

- For low energy solids, such as many polymers, the surface tension can be obtained from contact angle measurements. This was described in Section 7.4 and the exercise "Surface energy of polymers".
- From the value of the liquid: We measure the surface tension of the melt and rely on the fact that, close to the melting point, the free surface enthalpy of the solid is 10–20% larger.

- From the decrease of the lattice constant in small crystals caused by the compression due to the surface stress: The lattice spacings can be measured with the help of X-ray diffraction or by LEED experiments (see Section 8.7.2).

- From the work which is necessary for cleavage: We measure the work required to split a solid. The problem is that often mechanical deformations consume most of the energy and that the surfaces can reconstruct after cleaving [324].

- From adsorption studies with the help of inverse gas chromatography: The interaction between a stationary phase and a well known gas or solvent is measured as in standard gas chromatography. The results are, however, used to characterize the stationary phase. Therefore, let us assume that $n$ moles adsorb on a surface $A$. The interface excess concentration is $\Gamma = n/A$. With the Gibbs equation for a two-component system (Eq. 3.55) $\Gamma = -1/RT \cdot d\gamma/d(\ln P)$ we get

$$\gamma_S - \gamma_{SV} = -\int d\gamma = RT \cdot \int_0^{\ln P} \Gamma \cdot d(\ln P') = \frac{RT}{A} \cdot \int_0^{\ln P} n \cdot d(\ln P') \quad (8.7)$$

  We measure $n$ versus $P$, solve the integral, and get the difference between the surface tension of the bare solid surface and that of the surface–vapor interface $\gamma_S - \gamma_{SV}$.

- From mechanical measurement of the surface tension: If the surface tension of a solid changes, the surface tends to shrink (if $\Upsilon$ increases) or expand (if $\Upsilon$ decreases). This leads, for instance, to the deflection of bimetallic cantilevers or a contraction of ribbons.

- From the heat generated upon immersion: Material in the form of a powder with an known overall surface is immersed in a liquid. The free surface enthalpy of the solid is set free as heat and can be measured with precise calorimeters.

### 8.4.3 Surface steps and defects

If a crystal is cut at a small angle $\vartheta$ relative to a low-index surface, the surface exhibits steps or ledges that separate low-index terraces . The average distance of steps with height $h$ is given by $d = h/\sin\vartheta$, where $\vartheta$ is the angle between the low-index surface and the surface. Such a type of surface is called vicinal. Figure 8.12 shows an example of a vicinal surface on a simple cubic crystal. If there is an additional tilt of the cutting plane, the ledges will have kinks (Fig. 8.12, bottom).

Real surfaces will always exhibit a certain number of defects at temperatures above 0 K (Fig. 8.13). This is true in spite of the fact that defects have a positive energy of formation compared to an ideal crystalline surface. What stabilizes these defects is the change in entropy connected with the induced disorder. Therefore, a certain average number of defects — that increases with temperature — will be present.

As a consequence, real surfaces will not exhibit such evenly sized terrace or evenly spaced kinks as suggested by Fig. 8.12. The terrace-ledge-kink (TLK) model [332] can provide a more realistic description of vicinal surfaces. The distribution of the terrace widths is calculated taking into account the entropic repulsion between ledges. The confinement of a ledge between two neighboring steps (that cannot be crossed) leads to a reduction of the number

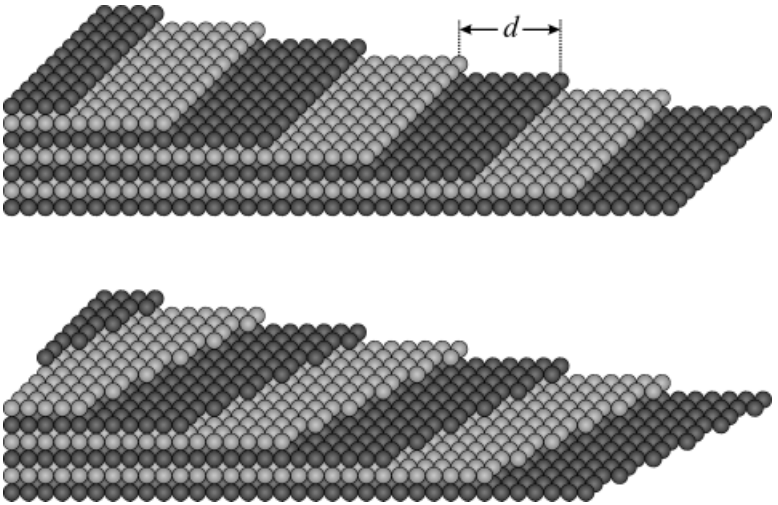



**Figure 8.12:** Top: Vicinal surface of cubic crystal that has been cut at a small angle relative to the {100} plane exhibiting monoatomic steps. Bottom: An additional tilt leads to kinks in the ledges.

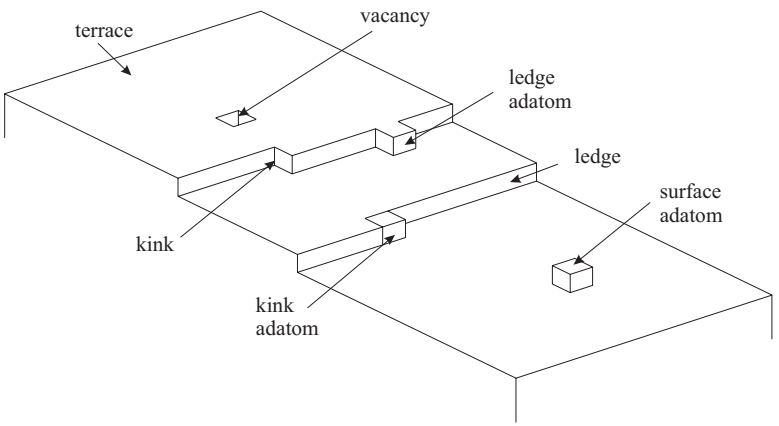



**Figure 8.13:** Schematic view of a vicinal surface with surface defects.

of possible kink configurations and therefore a decrease in entropy. This entropic repulsion follows a $d^{-2}$ law, where $d$ is the mean distance between the ledges [333]. Elastic strain and dipole interaction of steps may lead to an additional repulsion, which also follows a $d^{-2}$ dependence [334].

Defects lead to a roughening of crystal surfaces with increasing temperature as already predicted by Burton et al. [335]. However, calculations for low-index surfaces yield roughening transition temperatures well above the melting temperature. The reason is the high forma-

tion energy of a ledge on a low-index surface. For vicinal surfaces, where ledges are already present, roughening can occur by formation of kinks and roughening transition temperatures in the order of half the melting temperature are observed [336]. Hoogeman et al. [337] were able to directly image the roughening transition using high speed STM (see Section 8.6.3. From statistical analysis of the image data they were able to extract the kink creation energy and step interaction energy. An extensive review on the topic of steps on surfaces is Ref. [338].

An important class of defects are dislocations. Dislocations are not thermodynamically stable but kinetically trapped. The two primary types of dislocations are the edge dislocation and the screw dislocation (Fig. 8.14). An edge dislocation corresponds to an additional half-plane of atoms in the bulk crystal. Screw dislocations create a step on the surface that starts from the emergence of the dislocation at the surface.

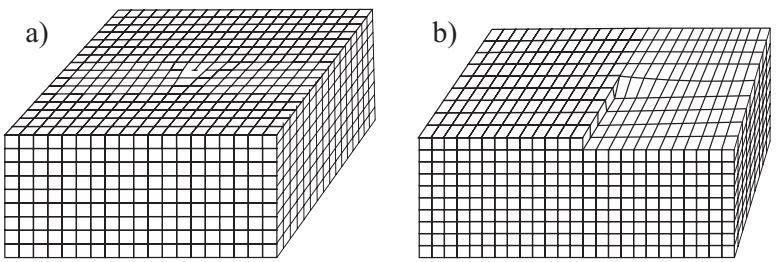



**Figure 8.14:** Examples of (a) an edge and (b) screw dislocation.

Dislocations are often characterized by their so-called Burgers vector. The Burgers vector of a dislocation is found in the following way: Draw a closed path consisting of an arbitrary sequence of lattice vectors within the perfect crystal lattice in the absence of any dislocations. Then follow the same sequence of lattice vectors while encircling the dislocation. This second path will not be closed due to the dislocation. The vector from the starting point to the end point of the second path is the Burgers vector.

■ **Example 8.4.** The burgers vector for the edge dislocation in Fig. 8.14 would be an in-plane vector perpendicular to the additional lattice plane, with a length of one lattice constant. For the screw dislocation, one gets a Burgers vector perpendicular to the surface. Its length is again identical to the lattice constant.

Surface defects always involve local variations in electronic states and binding energies. Therefore, surface defects are crucial in processes such as adsorption, nucleation, and surface reactions. For example, the step of a screw dislocation can eliminate the nucleation barrier for crystal growth.

## 8.5 Solid–solid boundaries

Until now we have concentrated on solid surfaces. This section is about solid–solid interfaces. Such interfaces play an essential role in the stability of materials. The stability of materials is a complex subject in itself and we will not deal with it within the scope of this book. We will only provide some basic concepts and classifications. In the semiconductor industry,

the production of solid–solid interfaces with well defined microscopic structure is of strong commercial interest.

There are different criterion of how to classify solid–solid interfaces. One is the sharpness of the boundary. It could be abrupt on an atomic scale as, for example, in III-IV semiconductor heterostructures prepared by molecular beam epitaxy. In contrast, interdiffusion can create broad transitions. Surface reactions can lead to the formation of a thin layer of a new compound. The interfacial structure and composition will therefore depend on temperature, diffusion coefficient, miscibility, and reactivity of the components. Another criterion is the crystallinity of the interface. The interface may be crystalline–crystalline, crystalline–amorphous, or completely amorphous. Even when both solids are crystalline, the interface may be disturbed and exhibit a high density of defects.

Thermodynamic considerations may help to find the most likely equilibrium shapes, but these may not be realized due to kinetic limitations. As a rule of thumb, minimizing interfacial energy usually means maximizing atomic matching to reduce the number of broken bonds.

For crystalline–crystalline interfaces we further discriminate between homophase and heterophase interfaces. At a homophase interface, composition and lattice type are identical on both sides, only the relative orientation of the lattices differ. At a heterophase interface two phases with different composition or/and Bravias lattice structure meet. Heterophase interfaces are further classified according to the degree of atomic matching. If the atomic lattice is continuous across the interface, we talk about a fully coherent interface. At a semicoherent interface, the lattices only partially fit. This is compensated for by periodic dislocations. At an incoherent interface there is no matching of lattice structure across the interface.

The most important example of homophase interfaces are grain boundaries. Figure 8.15 shows a model for a tilt grain boundary for a simple cubic crystal. The two grains are tilted relative to each other and joined to form the grain boundary. The former surface ledges become edge dislocations. The number of edge dislocations per unit length is given by:

$$\frac{1}{D} = \frac{2\sin(\vartheta/2)}{b} \tag{8.8}$$

for small angles $\vartheta$. Here, $b$ is the absolute value of the Burgers vector characterizing the dislocations (in this simple case, it is identical to the step height $h$). Each dislocation has an elastic strain energy that depends on the spacing of the dislocations. For small angle ($\vartheta \leq 15°$) grain boundaries, the grain boundary energy $\gamma_{GB}$ is calculated by summing up the contributions of all edge dislocations [340]. This results in the Read–Shockley equation

$$\gamma_{GB} = E_0\vartheta\left(1 + \ln\frac{b}{2\pi r_0} - \ln\vartheta\right) \qquad \text{with} \qquad E_0 = \frac{\tau b}{4\pi(1-\nu)} \tag{8.9}$$

Here, $\tau$ and $\nu$ are the shear modulus and Poisson ratio of the bulk material, and $r_0$ is the so-called core radius of the dislocation, that is a measure of the energy of a single dislocation ($\vartheta$ has to be given in radians to be dimensionless).

It is instructive to have a closer look at the dependence of $\gamma_{GB}$ on $\vartheta$ for small angles: $d\gamma_{GB}/d\vartheta$ goes to infinity for $\vartheta \to 0$. This implies that the grain boundary energy rises steeply with the introduction of the first ledge. The reason is the long-range stress field associated with an isolated dislocation. This is quite different to the solid–vapor interface, where the energy of a ledge is localized and no long range stresses occur. With increasing $\vartheta$, the slope becomes more shallow, as the dislocations get closer and their stress fields starts to cancel.

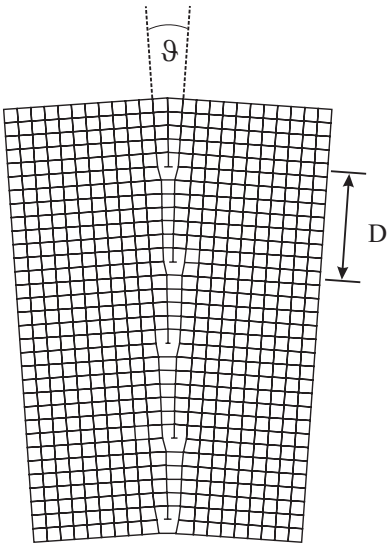



**Figure 8.15:** Symmetric small angle tilt boundary (drawn after [339]).

Another simple type of grain boundary is the twist boundary, where the lattice planes of the grains are rotated relative to each other. In this case the interface consists of a cross grid of screw dislocations. In the more general case, combinations of these two simple types of dislocation will occur.

For higher angle grain boundaries, the cores of the dislocations start to overlap and the simple summation of elastic strains no longer holds. To describe high angle grain boundaries, the coincident site lattice (CSL) model [341, 342] can be used. It starts by considering that if two contacting lattices are rotated relative to each other around a common lattice point, there will be coincidence of other lattice points at certain angles. The coincident points form a lattice themselves, which is termed a coincident site lattice. The degree of coincidence can be characterized by the parameter

$$\Sigma = \frac{\text{area of coincidence lattice cell}}{\text{area of original lattice cell}} \tag{8.10}$$

The grain boundary energy $\gamma_{GB}$ should be proportional to $\Sigma$. For small values of $\Sigma$ high coincidence occurs and the number of broken bonds can be minimized. $\Sigma = 1$ corresponds to complete coincidence of the ideal crystal. Experimentally it was found that the correlation between $\gamma_{GB}$ and $\Sigma$ is not that simple due to volume expansions or translations at the grain boundaries. A principal problem of the coincident site lattice model is that, even arbitrarily small variations of the lattice orientation lead mathematically to a complete loss of coincidence. This is physically not reasonable because an arbitrarily small deviation should have a small effect. This problem was solved by the O-lattice theory [343]. For a comprehensive treatment of solid–solid interfaces and grain boundaries, see Refs. [344, 345].

## 8.6 Microscopy of solid surfaces

In the next sections we introduce some of the most important experimental techniques of surface characterization. For the interested reader, a broad range of books on this topic is available (e.g. Refs. [346,347]). We start by discussing microscopy, continue with diffraction, and finally focus on spectroscopic methods.

### 8.6.1 Optical microscopy

Optical microscopy is often the first step in surface analysis, since it is fast and easy to perform. It can be an aid in selecting the area of interest on a sample for further analysis with more complex methods. The application of classical optical microscopy to surface science is, however, limited because the maximum lateral resolution is in the order of the optical wavelength ($\approx 500$ nm). For opaque solids, the light penetrates into the material, giving optical microscopy a poor surface sensitivity. In addition, the depth of field is limited which calls for flat, polished surfaces or allows only plane sections of the sample to be viewed.

Classical microscopy techniques are based on the incoherent scattering of visible light by the sample surface. Using coherent scattering of light, new methods of phase contrast microscopy became possible. It can be used to detect local changes in refractive index of transparent samples. Another method is interference microscopy where the light that is scattered from the surface is made to interfere with light that is reflected from an "ideally flat" surface within the instrument. In this way subnanometer resolution, normal to the surface, is achieved. The lateral resolution is still limited by the wavelength of light.

Confocal microscopy is a new technique in which the resolution of the optical microscope is increased and the background light is minimized. Rather than illuminating the whole sample only a tiny spot is illuminated and only the light coming from this spot is viewed. Focussing the illuminating light and viewing the scattered light is done via the same objective lens. The sample is raster-scanned in three dimensions and the light intensity is displayed versus the position on the sample. Three-dimensional images with a resolution of 200–400 nm laterally and 1–2 $\mu$m normal to the surface can be obtained.

Apart from optical microscopy, there are some other optical techniques which are truly surface sensitive and have found widespread use. Examples are ellipsometry (see Section 9.4.1), total internal reflection fluorescence (TIRF) [316], and surface plasmon resonance techniques [348].

### 8.6.2 Electron microscopy

With the development of quantum mechanics at the beginning of the 20th century it became clear that microscopic particles such as atoms, electrons and neutrons, in some cases behave like waves. Both views, the classical picture of a particle characterized by a momentum $p$, and that of a wave with a wavelength $\lambda$, are only two different, complementary, viewpoints of the same physical object. Both quantities are related by the de Broglie[2] relation $\lambda = h/p$, where

[2] Louis Victor de Broglie, 1892–1987. French physicist who proposed Eq. (8.11) in his Ph.D. thesis. Nobel prize in physics, 1929.

$h$ is the Planck constant. Some experiments are better understood in terms of the particle approach, others are better described in the wave picture.

To understand the resolution of electron microscopes and later diffraction techniques the wave approach is more instructive. The resolving power of electron microscopes is substantially better than that of a light microscope, since the wavelength of the electrons, given by the de Broglie relation

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2m_e E_{kin}}} \tag{8.11}$$

is smaller. Here, $m_e$ the electron mass, and $E_{kin}$ is their kinetic energy. In an electron microscope the electrons get their kinetic energy by an applied electric potential of $1-400$ kV. This leads to wavelengths of $0.4-0.02$ Å. Electron beams can be deflected by electromagnetic fields. This allows the construction of electron optics by using electromagnetic lenses. A broad overview can be found in Ref. [349].

We distinguish two basic types of electron microscope: the **transmission electron microscope** (TEM) and the **scanning electron microscope** (SEM). In transmission electron microscopy, an electron beam is transmitted through the sample. All electrons have the same energy and thus the same wavelength because they are accelerated by the same voltage. The scattered electrons are projected onto a phosphorous screen or an imaging plate to obtain the image. Since the electrons have to pass through the whole sample, the transmission electron microscope is not a surface-sensitive method and we will not discuss it further here.

Scanning electron microscopes are routine instruments to get a view of the surface structure with resolutions of typically 1–20 nm. Apart from this excellent lateral resolution, the high depth of field is an advantage of scanning electron microscopes. It enables us to get a three-dimensional impression of the sample topography. In a scanning electron microscope (Fig. 8.16) the electron beam is emitted from a heated filament or a field emission tip. The electrons are accelerated by an electric potential in the order of 1–400 kV. A condensor lens projects the image of the source onto the condensor aperture. The beam is focused by an objective lens and raster-scanned by scanning coils over the sample. It is important to minimize the spot size of the primary electron beam, as it determines the resolution of the instrument. When the primary electrons hit the sample surface, they pass part of their energy to electrons in the sample, resulting in the emission of secondary electrons. These secondary electrons have low energies ($\approx$ 20 eV) and therefore, only those secondary electrons escape from the sample, which were close ($\approx$ 1 nm) to the surface of the sample. This makes scanning electron microscopy highly surface sensitive. The secondary electrons are collected by a detector and their intensity is displayed versus the position of the primary beam on the sample.

To avoid electric charging by the electron beam, the objects must either be conducting or they have to be covered with a thin metal layer. Furthermore, the sample is placed in vacuum, at $\approx 10^{-7}$ mbar, so that the electrons are not scattered by gas molecules. In the new environmental scanning electron microscopes (ESEM) [350] it is also possible to get an image at gas pressures of up to 30 mbar. The trick with this type of instrument is to separate the sample chamber with the detector from the rest of the instrument by pressure-limiting apertures. These allow the electron beam to pass from the high-vacuum part with the electron gun and most of the electron pathway into the sample chamber while reducing the number of gas molecules that can enter the high-vacuum area. The ESEM uses a different

type of detector that makes use of the presence of the higher residual gas pressure in the sample chamber. The secondary electrons emitted from the sample are accelerated towards the positively biased detector. Collisions between the electrons and the residual gas atoms set free additional electrons by ionization of the atoms. On one hand this leads to an amplification of the electron signal. On the other hand the remaining positively charged gas ions can neutralize the excess negative charge that would otherwise build up at the surface. This allows the investigation of non-conducting samples without the need for metalization.

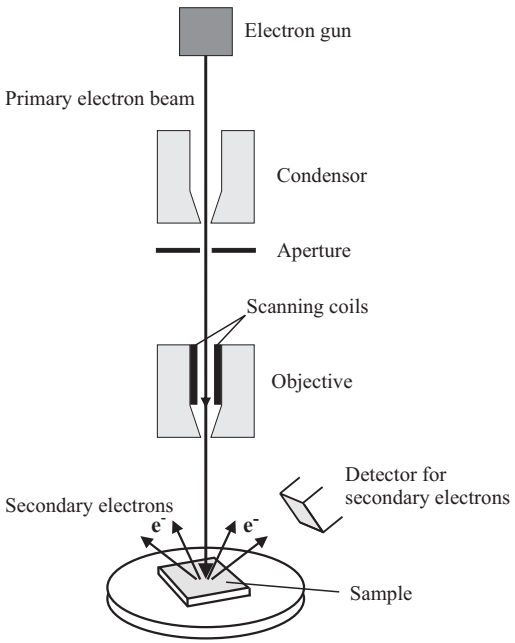



**Figure 8.16:** Schematic of a scanning electron microscope (SEM).

Depending on the type of scanning electron microscope, it is possible to obtain more information than just the sample topography. Besides the secondary electrons that are detected for the topographic imaging, there are also elastically backscattered primary electrons and X-rays generated by the interaction of the primary electrons with the sample atoms. The backscattering probability depends on the mean atomic number of the material. This imaging mode allows material contrast but is less surface sensitive since the penetration depth of the backscattered electrons is in the range of some 100 nm. The X-ray emission can be used for X-ray spectroscopy of the sample that allows quantitative analysis of the sample composition (see description of EDX below).

### 8.6.3 Scanning probe microscopy

Scanning probe microscopes (SPMs) have dramatically increased our possibilities for analyzing surfaces. The two most important representatives from the family of scanning probe

microscopes are the **scanning tunneling microscope** (STM) and the atomic force microscope (AFM) also called the scanning force microscope (SFM). In the scanning tunneling microscope [352] a fine metal tip is brought into close proximity of the sample. The sample must be electrically conductive. Between the sample and tip a voltage $V$ of typically 0.1–1 V is applied. At a distance of approximately 1 nm, a tunnel current begins to flow. The current intensity depends strongly on the distance. The sample is scanned underneath the metal tip by a scanner which consists of piezoelectric crystals. Usually the tunneling current is kept constant and an electronic feedback system regulates the $z$ (height) position of the sample while scanning. Practically in this way the distance $D$ is kept constant. This height information of the sample is finally plotted in the *xy*-plane. In this way a picture of the sample surface is obtained.

The tunnel current $I$ is proportional to

$$I \propto \frac{V\sqrt{\Phi}}{D} \cdot e^{-KD\sqrt{\Phi}} \tag{8.12}$$

Here, $\Phi$ is the mean work function of tip and sample, $\Phi = 1/2 \cdot (\Phi_{tip} + \Phi_{sample})$ and $K$ is a constant. In addition, the current is approximately proportional to the density of electronic states near the Fermi level [353]. With the scanning tunneling microscope, the density of the electron states is thus scanned.

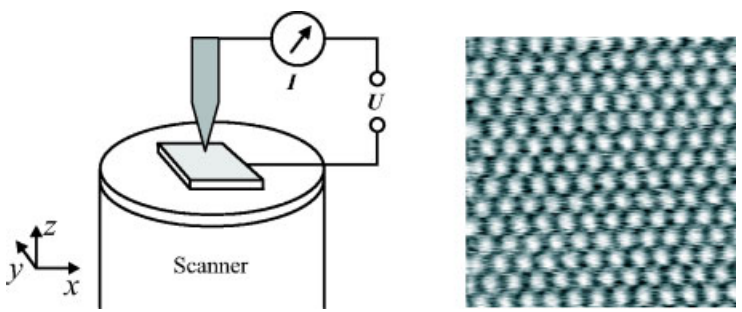



**Figure 8.17:** Left: Schematic of a scanning tunneling microscope (STM). Right: STM image ($2.7 \times 2.7$ nm) of the atomic structure of a copper (111) surface imaged in an aqueous medium after electrochemical cleaning [357]. The image was kindly provided by P. Broekmann and K. Wandelt.

Atomic resolution up to the imaging of atomic defects in crystal surfaces, is possible with the scanning tunneling microscope (Fig. 8.17). Surfaces can be scanned in vacuum, air and, with a little more effort and experience, in liquid. With the scanning tunneling microscope we are, however, restricted to electrically conductive samples such as metals or semiconductors. In addition, it is possible to scan thin non-conducting layers ($\approx$ 1 nm thick) on conducting substrates. Thus, it was possible to take images of certain organic monolayers or adsorbates.

One might think that it is difficult to produce tips which allow atomic resolution imaging. This is fortunately not the case. Often a wire made of Pt/Ir which is just cut with scissors produces atomically resolved images. Some researchers further etch the tip to get more reproducible results.

The restriction to conducting samples was overcome with the invention of the **atomic force microscope** (AFM) [143]. In this microscope, a tip attached to a cantilever is scanned over the sample surface (Fig. 8.18). The end of the tip is usually in mechanical contact with the sample. The sample is moved by a piezoelectric scanner in the xy-plane, as with the scanning tunneling microscope. The critical element in an atomic force microscope is the cantilever with the tip. Both are usually made using microfabrication techniques. In order to measure the up and down movements of the tip, a laser beam is focused on the back of the cantilever. From there the laser beam is reflected towards a photodetector. If the cantilever moves, the position of the reflected beam changes. With the help of the photodetector this change is converted into an electrical signal. As in the scanning tunneling microscope an electronic feedback control keeps the cantilever deflection constant by adjusting the $z$ position of the sample. This height information of the sample is finally plotted in the $xy$-plane and a topographic picture of the sample surface is obtained.

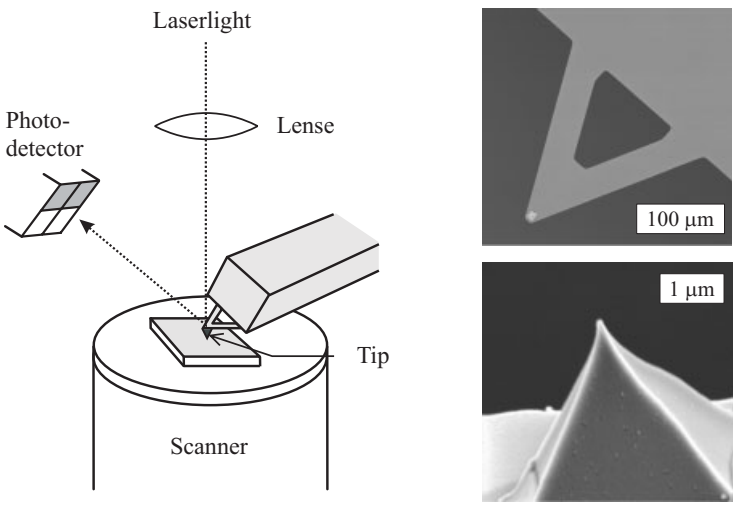



**Figure 8.18:** Schematic of an atomic force microscope (AFM). On the right, scanning electron micrographs show a cantilever (top) and a tip (bottom) in more detail. The tip, which in operation points downwards to the sample, is pointing towards the observer (top) and upwards (bottom).

The resolution of the atomic force microscope depends on the radius of curvature of the tip and its chemical condition. Solid crystal surfaces can often be imaged with atomic resolution. At this point, however, we need to specify what "Atomic resolution" is. Periodicities of atomic spacing are, in fact, reproduced. To resolve atomic defects is much more difficult and usually it is not achieved with the atomic force microscope. When it comes to steps and defects the scanning tunneling microscope has a higher resolution. On soft, deformable samples, e.g. on many biological materials, the resolution is reduced due to mechanical deformation. Practically, a real resolution of a few nm is achieved.

One danger in atomic force microscopy is the destruction of fragile objects on the sample. Although the force applied by the tip to the sample is only of the order of 1 nN, the pressure

can easily reach 1000 bar because the contact area is so small. A way to avoid deformation of fragile objects is to use the so-called tapping mode. In tapping mode the cantilever is vibrated at its resonance frequency. At the end of the cantilever, where the tip is, the vibration amplitude is typically 1–10 nm. When we approach the surface with such a vibrating cantilever, at some point the amplitude will decrease, simply because the tip starts to hit the surface. Instead of scanning the surface at constant deflection or constant height, we scan at constant reduction of the vibration amplitude. As a result, most of the time the tip is not in actual contact with the surface (Fig. 8.19). It just hits the surface for a short time at intervals given by the resonance frequency. The tapping mode is often less destructive than the contact mode because the contact time is very short and shear is prevented. In addition we get information about the local mechanical properties such as the elasticity [354]. One disadvantage of the tapping mode is the slightly lower resolution. It is usually impossible to image crystals with atomic resolution.

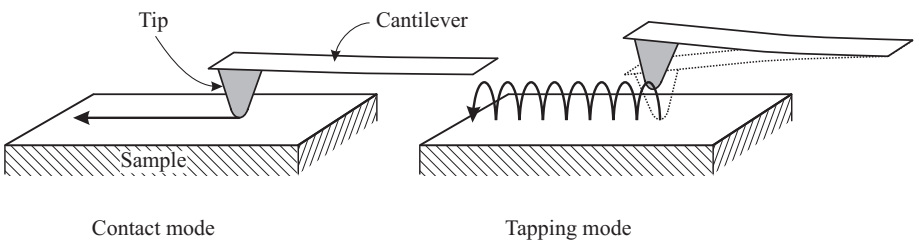



**Figure 8.19:** Actual movement of the AFM tip with respect to the sample in the standard contact mode and in the tapping mode.

With the atomic force microscope surfaces can be scanned in UHV, air, and most favorably in liquids. Liquids have the advantage that the force between tip and sample, which might lead to a possible deformation of fragile sample structures, is smaller than in air or UHV.

During the last two decades a wealth of different scanning probe microscope techniques based on scanning tunneling and atomic force microscopy has evolved, that allows us to gain more information than just topography. Examples are friction, adhesion, elasticity, conductance, electron densities, magnetization, and surface charges. For more information we refer to Refs. [355, 356].

## 8.7 Diffraction methods

For solid surfaces with crystalline structure, we can apply diffraction techniques for analysis. In a diffraction experiment the sample surface is irradiated with electrons, neutrons, atoms, or X-rays and the angular distribution of the outgoing intensity is detected. The analysis of diffraction patterns is a formidable task and in the first subsection we only introduce a simple case, which, nevertheless, contains the main features. A more general formalism for the interested reader is described in the Appendix.

### 8.7.1 Diffraction patterns of two-dimensional periodic structures

To analyse diffraction patterns we use the wave picture. The incident beam is treated as a coherent wave with a large extension perpendicular to its direction of propagation. "Large" means that it extends over a much larger distance than a typical atomic spacing. We start our discussion by asking: What is the diffraction pattern of a row of atoms of equal spacing $d$, which is irradiated perpendicular to its axis (Fig. 8.20)? The atoms are supposed to partially scatter the incoming beam. They act as new sources of spherical waves which are all in phase. We observe the resulting diffraction pattern far away from the sample as compared with the size of the sample. This implies that the rays at a certain point of observation are almost parallel, although they emanate from different atoms of the sample.

Rays coming from two neighboring atoms interfere constructively if the difference in path-length is just an integer number of wavelengths. This is the case for

$$\Delta = n\lambda = d \sin \vartheta, \tag{8.13}$$

where $n$ is an integer that is called the order of the diffraction peak and $\vartheta$ is the angle of observation. This is fulfilled for all outgoing waves that are on a cone around the axis along the one-dimensional structure, with a cone angle of $2 \times (90° - \vartheta)$. If we insert a detector plate, the observed diffraction pattern would consist of hyperbolic curves.

A real two-dimensional crystal consists of a periodic structure in not only one, but in two directions. This second periodic structure in another direction also leads to a condition for constructive interference. The conditions for both directions have to be fulfilled at the same time. Both conditions are only fulfilled where the cones of both orientations intersect. As a result constructive interference occurs only on lines starting from the point of incidence. On a detector plate we would observe spots where these lines cross the plate.

### 8.7.2 Diffraction with electrons, X-rays, and atoms

Surface diffraction methods should fulfill three conditions:

1. The resolution must lie in the range of some atomic layers, since only over this depth the surface differs from the volume. This criterion together with the angle of incidence sets an upper limit for the wavelength.

2. The method must be surface sensitive, i.e. the observation depth must substantially be defined by the first atomic layer. For this reason the penetration depth of the beam should be small.

3. The method must be nondestructive, i.e. no irreversible changes in the surface structure should occur.

Surface diffraction experiments have to be done in UHV. Otherwise the surfaces are covered with a monolayer of adsorbed molecules. At this point the reader might ask: why do we not have to use UHV in scanning tunneling or the atomic force microscope? In both techniques the tip penetrates through the surface contamination layer. In the scanning tunneling microscope it is often just invisible because contamination layers are usually not good conductors. In

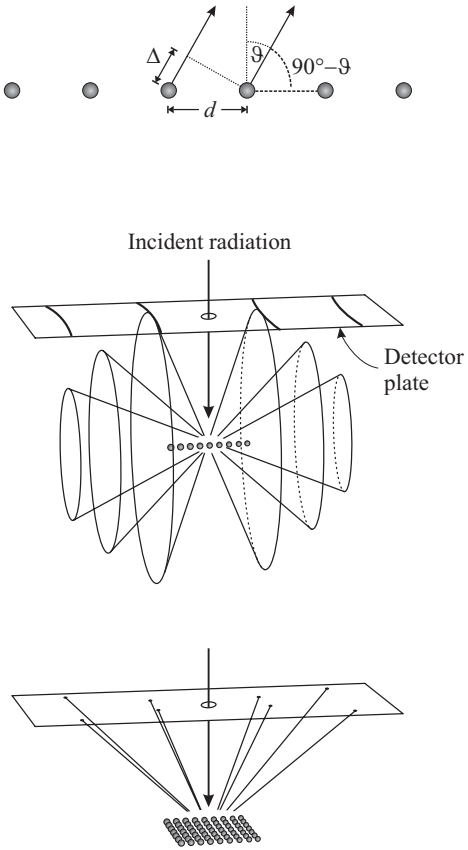



**Figure 8.20:** Condition for constructive interference for a one-dimensional array of scattering centers (top). Constructive interference occurs along the cones that reflect the rotational symmetry of the one-dimensional arrangement (middle). For a two-dimensional crystal, constructive interference is obtained along lines (bottom).

atomic force microscopy it might disturb the imaging process but often the tip mechanically pushes through.

Most commonly **low-energy electron diffraction** (LEED) is used for the determination of surface structures. With typical electron energies of 20–500 eV the de Broglie wavelength is at 3–0.5 Å. The penetration depth is relatively independent of the material and typically 4–10 Å. In a LEED experiment a well-collimated monoenergetic beam of electrons is directed towards the sample (fig. 8.21). The elastically backscattered electrons interfere and form a pattern of maxima and minima which is visualized on a fluorescent screen after amplification. Since the wavelength of the electrons lies in the range of atomic spacings, only few reflexes appear on the screen. If we increase the electron energy, the angle between the reflexes becomes smaller and the reflexes will move closer on the screen.

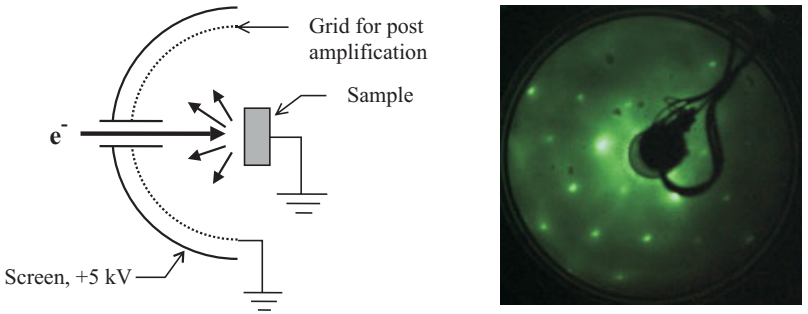



**Figure 8.21:** Schematic of a LEED setup (left) and a LEED diffraction pattern obtained from a Pt(111) surface with electrons of 350 eV energy (right). The image of the diffraction pattern was kindly provided by M. Zharnikov and M. Grunze.

■ **Example 8.5.** With an electron energy $E$ of 50 eV, the wavelength is $\lambda = h/\sqrt{2m_eE} = 1.73$ Å. With a lattice constant of 3.0 Å the first reflex appears under an angle of $35.2°$. If we increase the electron energy to 100 eV or 200 eV, then the wavelengths are $\lambda = 1.23$ Å and 0.87 Å. The corresponding first intensity maxima appear under angles of $24.2°$ and $16.9°$, respectively.

The incident electrons might change the sample structure. For low-energy electrons these might be electronic excitations, which quickly disappear in metals and semiconductors, which can however, be severe in ionic or covalent crystals. The more insulating a solid is, the higher is the danger of changes such as desorption, place change, dissociation, and other secondary processes, which are caused by charge transfer in atoms.

High-energy electrons with energies in the range of 1–5 keV can penetrate deeper into a solid. Therefore in **high-energy electron diffraction**, also called reflective high energy electron diffraction, RHEED, HEED, or RED, we usually work at a grazing incidence angle, in order to achieve a high surface sensitivity with small penetration depth. The technology can usually not compete with LEED and is used if there is not enough space in the UHV chamber to accommodate LEED.

The interaction of X-rays with matter is weak compared to that of electrons. This leads to a large penetration depth of some $\mu$m

and allows surface inspection only at grazing incidence. The technique is called **grazing incidence X-ray diffraction** (see also Section 13.3.2). At the same time the diffraction intensity from a single surface monolayer is small and particularly strong X-ray sources are necessary, such as synchrotrons. X-ray diffraction has a big advantage: Since interactions with the sample are scarce, multiple scattering effects can often be neglected which makes a comparison with calculations much simpler.

Atomic beams are suitable for diffraction experiments. The repulsive interaction with the surface is so strong that low-energy atoms are already reflected by the topmost atomic layer. Due to its pronounced sensitivity to the topmost surface layer, the method of **atomic beam diffraction** is especially useful for the study of adsorbates and superlattices. Typical energies of the incident atoms are under 0.1 eV. At such low energies no radiation damage occurs. In

order to generate a monochromatic atom beam the atoms are allowed to escape from a gas reservoir through a nozzle. Apertures and a chopper collimate and monochromatize the beam.

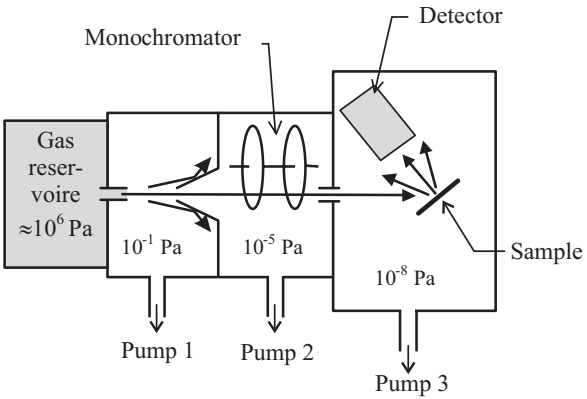



**Figure 8.22:** Schematic of an atomic beam diffractometer. Pressure-limiting apertures separate the zones with different vacuum quality.

## 8.8 Spectroscopic methods

In the previous chapter, we introduced methods to analyze the structure of surfaces. Now we discuss methods to analyze the chemical composition and the electronic energies of molecules at the surface.

### 8.8.1 Spectroscopy using mainly inner electrons

Spectroscopy of the inner electrons tells us something about the elemental composition of the surface atoms because the inner electrons are only weakly affected by chemical bonds. In **X-ray photoemission spectroscopy**, (XPS or ESCA for "electron spectroscopy for chemical analysis") the sample is exposed to monochromatic photons of high energy ($> 1$ keV). The photons excite sample atoms and the kinetic energy of emitted photoelectrons is measured. The kinetic energy is $h\nu - E_b$, where $h\nu$ is the energy of the photon and $E_b$ is the binding energy of the electron.

■ **Example 8.6.** Fig. 8.23 shows two X-ray photoemission spectra. On the left, a survey spectrum recorded on palladium which was radiated by Mg $K_\alpha$ X-rays (1253.6 eV). The horizontal axis is inverted because, practically, the kinetic energy of the emitted electrons is plotted. The spectrum shows a small peak at 51 eV which corresponds to the binding energy of a 4P electron. A peak at 86 eV (4S) is not visible because the absorption cross-section is too low. At 335 eV, a peak of the 3D electron is visible followed by two peaks corresponding to two 3P orbitals (at 531 and 559 eV). The highest binding energy is observed for the 3S electrons (670 eV). On the right, a detail of a nitrogen spectrum is

shown. It was recorded on a mixed monolayer of dipalmitoyl phospatidylcholine (DPPC) and the lipopeptide surfactin (at 1:1 molar ratio) on mica deposited by Langmuir–Blodgett transfer at a film pressure of 20 mN/m (see Chapter 13). For XPS the film was irradiated with monochromatized Al $K_\alpha$ X-ray radiation (energy 1486.6 eV). The energy of electrons emitted from the nitrogen 1S orbital ($E_b = 399$ eV) was detected.

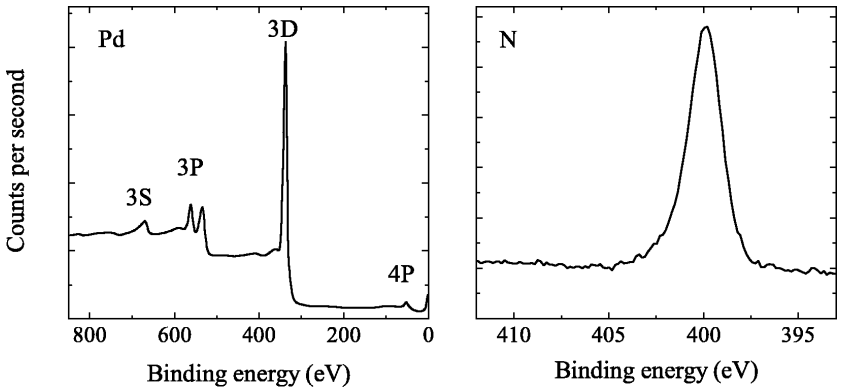



**Figure 8.23:** Left: XPS survey spectrum recorded on palladium. Right: Nitrogen 1S spectrum of a surfactin/DPPC monolayer on mica. The vertical scale shows the number of counts per second. Adapted from ref. [358].

Why is this technique surface-sensitive, although the incoming X-rays penetrate, typically, microns into the sample? It is surface sensitive because only those electrons which are emitted by atoms close to the surface have a chance to leave the sample and are detected.

With the availability of synchrotrons as high intensity tunable X-ray sources, another technique had a break-through. This technique is called **extended x-ray absorption fine structure** (EXAFS). In EXAFS the precise shape of an X-ray absorption band is measured. The photoelectron ejected from the absorbing atom can be described as an outgoing wave, that will be scattered by the surrounding atoms. The interference between the outgoing wave and the scattered waves leads to a fine structure of the absorption band. The periodicity of the fine structure is related to the distances between the absorbing atom and its neighbors. The amplitudes of the periodic peaks contain information about the number of neighboring atoms. Therefore EXAFS probes the local structure around the absorbing atoms. The advantage is that not only can crystals be analyzed, but also polycrystalline material without long-range order. This was, for instance, essential in the study of catalysis.

If an electron of an inner shell is removed by X-rays or primary electrons, an electron from an outer shell fills the hole. During this process either a photon with an energy that corresponds to the difference between the two involved energy levels is emitted or this energy is transferred to another outer electron. This outer electron is then emitted. It is called the Auger electron. We can either analyze the energy of the emitted X-rays, which is called **energy dispersive X-ray analysis** (EDX), or the energy of the emitted Auger electrons, which is called **Auger electron spectroscopy** (AES). The emission of an Auger electron is thus characterized by three energy levels. AES and EDX are suitable, like XPS, for an element

analysis. AES is more suitable for elements of low atomic numbers, while with EDX more heavier elements can be sensitively detected. EDX is often used for elemental analysis in electron microscopy, where the electron beam is used to eject the inner electrons. The three methods are schematically summarized in Fig. 8.24.

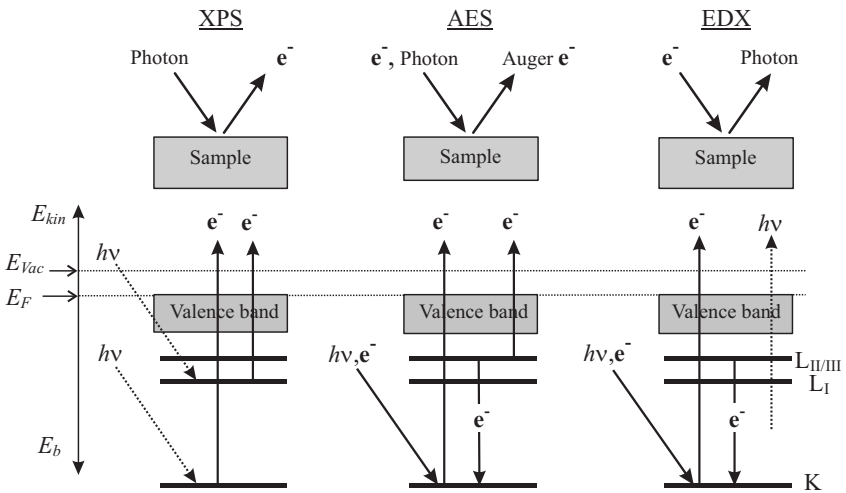



**Figure 8.24:** Term diagram for XPS, AES, and EDX. The vacuum energy $E_{vac}$ defines the zero point of the energy scale. The binding energy of electrons $E_b$, the Fermi energy $E_F$ and the kinetic energy of free electrons $E_{kin}$ are indicated.

Why is there no method, in which we excite with photons and detect the emitted photons? The reason is simply that such a method would not be surface sensitive. The incoming photons penetrate deeply into the solid and the emitted photons are still detected from large depths.

### 8.8.2 Spectroscopy with outer electrons

To find something out about molecules and bonds between atoms, the energies of the exciting photons or electrons have to be reduced. In contrast to the inner electrons which occupy isolated atomic orbitals, the energies of the outer electronic structure are determined by valence or conduction bands.

An important technique is **UV photoemission spectroscopy** (UPS) which is based on the outer photoelectric effect (in contrast to XPS, where we use the inner photoelectric effect). Photons with energies of 10–100 eV are used to ionize atoms and molecules at the surface. The energy of emitted electrons is detected. To study adsorption of molecules to surfaces, often difference spectra are analyzed which are measured before and after the adsorption. These difference spectra are compared to the spectrum of the molecules in the gaseous phase.

A versatile tool to analyze vibrations of surface atoms and adsorbed molecules is **high-resolution electron energy loss spectroscopy** (HREELS) [359]. Monoenergetic low-energy electrons (1–10 eV) are directed to the surface. Most of them are backscattered elastically.

Some of them, however, excite vibrations of surface atoms or molecules. Their energy is thus reduced by the energy of the vibration. This allows one to take energy spectra of surface excitations.

### 8.8.3 Secondary ion mass spectrometry

In secondary ion mass spectrometry (SIMS) the sample surface is sputtered by an ion beam and the emitted secondary ions are analyzed by a mass spectrometer (review: Ref. [360]). Due to the sputtering process, SIMS is a destructive method. Depending on the sputtering rate we discriminate static and dynamic SIMS. In static SIMS the primary ion dosis is kept below $10^{12}$ ions/cm$^2$ to ensure that, on average, every ion hits a "fresh" surface that has not yet been damaged by the impact of another ion. In dynamic SIMS, multiple layers of molecules are removed at typical sputter rates 0.5 to 5 nm/s. This implies a fast removal of the topmost layers of material but allows quantitative analysis of the elemental composition.

We start with a discussion of dynamic SIMS. In dynamic SIMS the most common primary ion sources are oxygen duoplasmatrons, that produce $O^-$ or $O_2^+$, or caesium guns, which provide $Cs^+$ ions. These two elements are preferred for their enhancement effects. Oxygen implantation leads to the formation of oxides in the surface layer. During the secondary ion emission process the bonds of the oxides break, mainly leaving the oxygen as a negative ion and the original surface material as a positive ion. Caesium acts as an efficient electron donor and mainly negative secondary ions are generated.

Only a small fraction of the sputtered material will actually be ionized. The ionization probability depends on the element species and on the matrix material.

Most SIMS instruments use magnetic sector mass analyzers, where the different mass-dependent deflection of the ions in a magnetic field is used to separate the elements. Another type are quadrupole mass analyzers. There the ions move in the center of a quadratic arrangement of four parallel cylindrical electrodes. A combination of static and alternating electric fields provide stable oscillations only for ions with a specific charge to mass ratio. Only these ions are able to exit the aperture at the end of the analyzer. After mass separation the ions have to be collected by a detector. An electron multiplier can be used to count single ions. This allows for high sensitivity but limits detectable count rates to $10^6$ ions per second. For higher ion currents, a Faraday cup is used to collect the ions and measure the accumulated charge.

It is also possible to use locally resolved SIMS. SIMS imaging allows one to visualize the lateral variation of secondary ion intensities. This can be achieved by two different methods:

- A focussed primary ion beam is raster-scanned over the surface and the intensity of the secondary ion current is stored as a function of beam position. The result is visualized on a computer screen. The lateral resolution is limited by the diameter of the primary ion beam that can be as small as 20 nm.

- In the ion microscope the sample is "illuminated" with a broad ion beam ($25 - 250\,\mu$m). The secondary ions are filtered by mass spectrometers that conserve their spatial distribution. Ions are then visualized by microchannel plates as image detectors. The maximum lateral resolution of this method is in the order of $1\,\mu$m.

The continuous sputtering of the surface during analysis allows for the measurement of depth profiles. Therefore, the secondary ion intensity of the element of interest is recorded versus

sputtering time. The time axis is then converted to a length scale by measuring the depth of the sputtering crater at the end of the experiment, e.g., using a profilometer.

For static SIMS, the time-of-flight SIMS (TOF-SIMS) technique is most suitable. For TOF-SIMS, a pulsed primary ion beam is used. The secondary ions emitted from the surface after one pulse, are accelerated by an electrical potential $U$. These ions gain the velocity $v$ that depends on the mass $m$ and charge $Q$ of the ion according to:

$$v = \sqrt{\frac{2QU}{m}} \tag{8.14}$$

Therefore, the ions will hit the detector at distance $L$ after a time span

$$\Delta t = \frac{L}{v} = \frac{L}{\sqrt{2QU/m}} \tag{8.15}$$

In this way, the mass of the ions can be deduced from their arrival time at the detector and a complete mass spectrum is acquired for every pulse. Since after each pulse we have to wait until all ions have arrived at the detector, TOF-SIMS is limited to low sputter rates. This does not allow depth profiling, but makes TOF-SIMS virtually non-destructive compared to dynamic SIMS.

## 8.9 Summary

- Crystalline surfaces can be classified using the five two-dimensional Bravais lattices and a basis. Depending on the surfaces structure, the basis may include more than just the first surface layer. The substrate structure of a surface is given by the bulk structure of the material and the cutting plane. The surface structure may differ from the substrate structure due to surface relaxation or surface reconstruction. Adsorbates often form superlattices on top of the surface lattice.

- For the study of crystalline surfaces, ultrahigh vacuum (UHV) is required. The preparation of clean crystalline surfaces is usually carried out within the UHV system by cleavage, sputtering, evaporation, thermal treatment, or molecular beam epitaxy.

- The creation of a new surface on a solid can be done either plastically — described by the surface tension — or elastically — described by the surface stress. The equilibrium shape of a crystal can be determined from the Wulff construction. Surface energies can be calculated using a nearest-neighbor broken bond model: in first approximation, the surface energy corresponds to the energy required to break all bonds of the bulk crystal along the cutting plane. In practice, surfaces may lower their energy by surface relaxation or surface reconstruction. Real surfaces have defects. On vicinal surfaces, terraces, ledges and kinks are present. Other important types of defects are screw and edge dislocations.

- Grain boundaries are the most important type of homophase solid–solid interfaces. An example of crystalline heterophase interfaces with high relevance for technological applications are semiconductor heterostructures.

- Optical microscopy may be used for surface examination but in many cases it is not really surface sensitive. Scanning electron microscopy (SEM) allows to study surfaces with high resolution in vacuum. The scanning tunneling microscope (STM) can achieve atomic resolution on conducting substrates. The atomic force microscope (AFM) can be used on most solid surfaces for high resolution imaging. Diffraction techniques using X-rays, electrons, or atomic beams can reveal the surface structure of crystalline materials. Spectroscopic methods like XPS, EDX, AES, UPS, and SIMS can give information on the surface composition and the electronic energies at a surface.

## 8.10 Exercises

1. Draw schematically a hexagonal fcc(111) lattice and indicate the $(1 \times 2)$ and the $(\sqrt{7} \times \sqrt{7})\,\mathrm{R}\,19.1^\circ$ overlayer structures.

2. In the chapter about the AFM it is mentioned that the force between tip and sample is usually lower in liquids than in air or UHV. Why?

3. Calculate the structure factor of the diamond lattice (see Appendix).

4. You would like to analyze the atomic structure of Pd by grazing incidence X-ray diffraction. After penetrating a distance $x$, the intensity of the beam is decreased by $I = I_0 \cdot e^{-\mu x}$. Your diffractometer uses photons of 10 keV energy. Is the wavelength sufficiently small to analyze atomic structures? At this energy a handbook tells you that the photon attenuation coefficient, $\mu/\rho$, is 691 cm$^2$/g. The density is 12.0 g/cm$^3$. Assume you want 20% of the incident X-rays to be scattered within the topmost layer of 1 nm thickness. Which angle do you have to choose?

# 9 Adsorption

## 9.1 Introduction

In this chapter we deal with adsorption. Adsorption is the accumulation of a substance at an interface. We first concentrate on the simplest case: Adsorption of small, uncharged molecules to the solid–gas interface. Unless the solid is in ultrahigh vacuum, gas molecules will adsorb to its surface due to attractive intermolecular forces such as the omnipresent van der Waals forces. The amount adsorbed is determined by several parameters. The most important one is the partial pressure of the molecules $P$. At the surface, the rotational and vibrational freedom of adsorbed molecules is usually reduced. Even their electrical properties may change. Some molecules diffuse laterally or the molecules might react on the surface; these processes are highly important in the understanding of catalysis. Finally, molecules might desorb again into the gaseous phase. Adsorption and desorption rates determine the equilibrium amount on the surface.

We extend our description to adsorption at the solid–liquid interface. For many systems we can use the same models as for gas adsorption on a solid surface, we only have to replace the pressure $P$ by the concentration $c$. The adsorption of macromolecules to surfaces is briefly discussed in Section 10.3.2. For macromolecules desorption is often negligible and thermodynamic equilibrium is only reached after a very long time, if at all.

Instructive reviews about adsorption are Refs. [8,361]. A good overview on the adsorption of water to solid surfaces is Ref. [362].

### 9.1.1 Definitions

Let us first introduce the most important definitions. The material in the adsorbed state is called **adsorbate**. The substance to be adsorbed (before it is on the surface) is called the **adsorpt** or **adsorptive**. The substance, onto which adsorption takes place, is the **adsorbent** (Fig. 9.1).

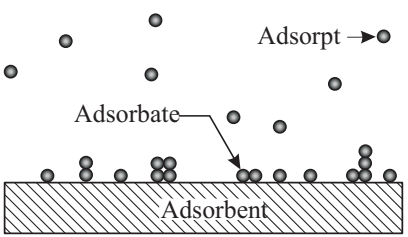



**Figure 9.1:** Definitions of adsorbent, adsorpt, and adsorbate.

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

An important question is how much of a material is adsorbed to an interface. This is described by the adsorption function $\Gamma = f(P, T)$, which is determined experimentally. It indicates the number of adsorbed moles per unit area. In general, it depends on the temperature. A graph of $\Gamma$ versus $P$ at constant temperature is called an **adsorption isotherm**. For a better understanding of adsorption and to predict the amount adsorbed, **adsorption isotherm equations** are derived. They depend on the specific theoretical model used. For some complicated models the equation might not even be an analytical expression.

All gases adsorb to solid surfaces below a critical temperature because of the van der Waals attraction. In general, when adsorption is dominated by physical interaction rather than chemical bonding, we talk about **physisorption**. Physisorption is characterized by several features:

- The sublimation energy is in the order of 20–40 kJ/mol.
- The adsorbate is still relatively free to diffuse on the surface and to rotate.
- The molecular structure of the solid does not change with physisorption except for some molecular solids (e.g. ice, paraffin, polymers).
- An adsorption equilibrium is quickly established. When lowering the pressure the gas desorbs reversibly (except in porous solids).

If the adsorption energy is of the order of chemical binding energies we talk about **chemisorption**. Characteristic properties are:

- Typical sublimation energies of 100–400 kJ/mol.
- Often there are specific binding sites. The adsorbate is relatively immobile and usually does not diffuse on the surface.
- Even on covalent or metallic solids there is often a surface reconstruction.
- Due to the strong binding, experiments in UHV are possible because the molecules practically do not desorb.

Oxidation can be viewed as the chemisorption of oxygen. For example, nickel and silicon are oxidized at ambient conditions. The resulting oxide layer is thermodynamically more stable and passivates the pure material below it. Another important example is the oxidation of aluminum which provides the metal with a very hard roughly 100 nm thick aluminum oxide ($Al_2O_3$) layer. To stabilize the aluminum surface even more and to passivate it against reactive chemicals the thickness of the oxide layer can be increased electrochemically. This procedure is called the eloxal process (*el*ectrolytical *ox*idation of *al*uminum).

### 9.1.2 The adsorption time

A useful parameter to characterize adsorption is the adsorption time. Let us first assume that no forces act between the surface and a gas molecule. Then, if a molecule hits the surface, it is reflected elastically with the same energy. An energy transfer between the surface and gas

molecule does not take place. As a consequence "hot" molecules do not cool down even when they hit a cold surface. The residence time in proximity to the surface can be estimated by

$$\tau = \frac{2\Delta x}{\bar{v}_x} \approx \frac{2\Delta x}{\sqrt{k_B T/m}}, \tag{9.1}$$

where $\Delta x$ is the thickness of the surface region and $\bar{v}_x$ is the mean velocity normal to the surface. Example: $N_2$ at 25°C, $\Delta x = 1$ Å, $\bar{v}_x \approx 300$ m/s, $\tau \approx 7 \times 10^{-13}$ s. This is in the order of typical vibration periods of $10^{-13}$ s.

An attractive force between the gas molecule and the surface increases the average residence time of the molecule at the surface to

$$\tau = \tau_0 \cdot e^{Q/k_B T}, \tag{9.2}$$

with $\tau_0 \approx 10^{-13}...10^{-12}$ s. $Q$ is the heat of adsorption. To be more precise we have to identify the inverse of $\tau_0$ with a surface bond vibration frequency. Values for various atoms and molecules adsorbed to well defined surfaces are given in Table 9.1. Heats of adsorption up to 10 kJ/mol refer to practically no adsorption and residence times are below 10 ps. $Q =$ 40 kJ/mol is characteristic for physisorption. Residence times become significantly longer and depending on the precise value of $Q$, can assume quite different values. Chemically adsorbed molecules ($Q \geq 100$ kJ/mol) practically do not leave the surface again.

**Table 9.1:** Heats of adsorption $Q$, surface bond vibration frequencies $\tau_0^{-1}$, and adsorption times $\tau$ at 27°C. Results from Ref. [363].

| | H/W(100) | Hg/Ni(100) | CO/Ni(111) | $N_2$/Ru(100) | Xe/W(111) |
|---|---|---|---|---|---|
| $Q$ (kJ mol$^{-1}$) | 268 | 115 | 125 | 31 | 40 |
| $\tau_0^{-1}$ (Hz) | $3 \times 10^{13}$ | $10^{12}$ | $8 \times 10^{15}$ | $10^{13}$ | $10^{15}$ |
| $\tau$ (s) | $10^{33}$ | $10^{8}$ | $7 \times 10^{5}$ | $3 \times 10^{-8}$ | $9 \times 10^{-9}$ |

Another useful parameter is the **accommodation coefficient** $\alpha$. The accommodation coefficient is defined by the temperature of the molecules before the impact $T_1$, the surface temperature $T_2$, and the temperature of the reflected molecules $T_3$ [364]:

$$\alpha = \frac{T_3 - T_1}{T_2 - T_1} \tag{9.3}$$

For an elastic reflection, the mean velocity of the molecules before and after hitting the surface are identical and so are the temperatures: $T_1 = T_3$. Then $\alpha = 0$. If the molecules reside a long time on the surface they have the same temperature, after desorption, as the surface: $T_2 = T_3$ and $\alpha = 1$. Thus, the accommodation coefficient is a measure of how much energy is exchanged before a molecule leaves the adsorbent again.

### 9.1.3 Classification of adsorption isotherms

Depending on the physicochemical conditions, a great variety of adsorption isotherms are experimentally observed. Eight common examples are shown in Fig. 9.2 [7, 365].

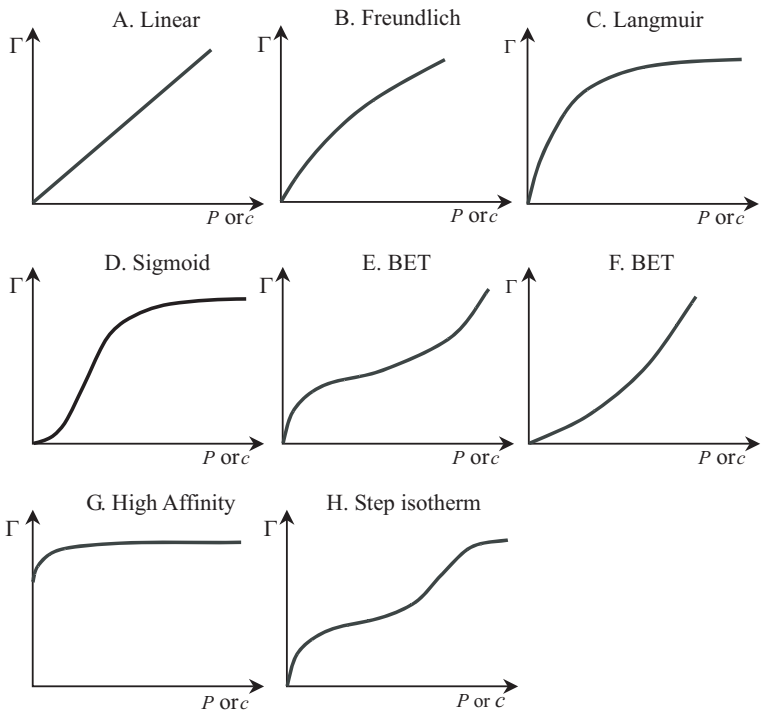



**Figure 9.2:** Schematic plot of eight types of adsorption isotherms commonly observed. If adsorption from the gas phase is studied, the abscissa is the partial pressure $P$. For adsorption from solution the concentration $c$ is used.

The most simple type, A, is that of a linear increase. It is described by the Henry adsorption isotherm equation:

$$\Gamma = K_H P \tag{9.4}$$

$K_H$ is a constant in units of mol m$^{-2}$ Pa$^{-1}$ for gases and L/m$^2$ for solutions. It is the ideal limiting law for low $\Gamma$.

Type B is very common. It is concave with respect to the abscissa. Most surfaces are heterogeneous. There are adsorption sites, which have a high affinity, and regions, which have a low affinity. The high affinity sites are occupied first, which accounts for the steep increase at low pressure. Another reason is sometimes a lateral repulsion between adsorbed molecules. This type of adsorption isotherm is described by the Freundlich[1] adsorption isotherm equation [366]:

$$\Gamma = K_F \cdot P^q \tag{9.5}$$

$K_F$ and $q$ ($q < 1$) are constants.

[1] Herbert Max Finlay Freundlich, 1880–1941. German physicochemist.

Type C is called the Langmuir[2] type because it can be described by the Langmuir adsorption isotherm equation:

$$\theta = \frac{K_L P}{1 + K_L P} \qquad \text{with} \qquad \theta = \frac{\Gamma}{\Gamma_{mon}} \tag{9.6}$$

Here, $\theta$ is the relative coverage and $K_L$ is a constant, called the "Langmuir constant". $\Gamma_{mon}$ is the maximum amount adsorbed which, in the case of Langmuir adsorption, is a monolayer. Type C adsorption isotherms are characterized by a saturation at high concentrations. A possible reason is that the surface is completely filled with adsorbed molecules. Langmuir adsorption is often observed for the adsorption from solution but only rarely for the adsorption of gases. This type of adsorption isotherm can also be observed for porous materials. Once all pores have been filled the isotherm saturates (see Section 9.4.3).

A sigmoidal isotherm (type D) indicates cooperative effects. A molecule binds to the surface better if it can interact with a neighboring adsorbed molecule. As a consequence of this lateral interaction two-dimensional condensation occurs. In order to observe sigmoidal isotherms, flat and homogeneous adsorbents are required.

Type E is common for the adsorption of gases. Usually the first concave part is attributed to the adsorption of a monolayer. For higher pressures more layers adsorb on top of the first one. Eventually, if the pressure reaches the saturation vapor pressure, condensation leads to macroscopically thick layers. It can be described by the BET adsorption isotherm equation Eq. (9.37) (see below).

Type F is expected if the binding of the first monolayer to an adsorbent is weaker than the binding of molecules to already adsorbed molecules. This is the case if the heat of adsorption is lower than the heat of condensation. It can also be described by the BET theory.

Type G is a high-affinity adsorption isotherm. The molecules bind so strongly that no rest can be detected in the solution or gas phase. The difference to the Langmuir type is quantitative not qualitative. It is often observed for polymer or proteins adsorbing from solution.

Step isotherms (type H) are observed with porous materials and characterized by a second inhibition. At low pressure a single layer of molecules adsorbs to the surface as for Langmuir adsorption. At intermediate pressures, multilayers start to form and the pores are filled. The saturation at high pressures is caused by the reduction of effective surface area once the pores have been filled.

### 9.1.4 Presentation of adsorption isotherms

An adsorption isotherm is a graph of the amount adsorbed versus the pressure of the vapor phase (or concentration in the case of adsorption from solution). The amounts adsorbed can be described by different variables. The first one is the surface excess $\Gamma$ in mol/m$^2$. We use the Gibbs convention (interfacial excess volume $V^\sigma = 0$). For a solid surface the Gibbs dividing plane is localized directly at the solid surface. Then we can convert the number of moles adsorbed $N^\sigma$ to the surface excess by

$$\Gamma = \frac{N^\sigma}{A}, \tag{9.7}$$

where $A$ is the total surface area.

[2] Irving Langmuir, 1881–1957. American physicist and chemist, spent most of his time at the General Electric Company. Nobel prize for chemistry in 1932.

Adsorption is often studied using powders or porous materials because the total surface area is large even for small amounts of adsorbent. In a typical experiment the volume ($V$) or the mass ($m = V/\rho$) adsorbed per gram of adsorbent, is measured. Theoretical models always describe an adsorption per surface area. In order to compare theoretical isotherms to experimentally determined adsorption results, the specific surface area needs to be known. The specific surface area $\Sigma$ (in $m^2$/kg) is the surface area per kg of adsorbent. Once the specific surface area is known, the area can be calculated by $A = m_{ad}\Sigma$, where $m_{ad}$ is the mass of the adsorbent.

The abscissa of an adsorption isotherm plot is usually the pressure. For gases it is given in Pa. If we are studying the adsorption of a vapor in equilibrium with its liquid the relative vapor pressure $P/P_0$ is plotted.

## 9.2 Thermodynamics of adsorption

### 9.2.1 Heats of adsorption

Heats of adsorption are important characteristics of adsorption because they provide information regarding the driving forces for adsorption. Several heats, energies, enthalpies, and other quantities of adsorption have been defined. For a detailed discussion see Refs. [367, 368].

We first introduce the **integral molar energy of adsorption**:

$$\Delta_{ad}U_m^{int} = U_m^{\sigma} - U_m^{g} \tag{9.8}$$

It is the energy difference between $N^{\sigma}$ moles of gas adsorbed $U_m^{\sigma}$ (per mol) and the same amount free in the gas phase $U_m^{g}$. The next important quantities are the **integral molar enthalpy of adsorption**

$$\Delta_{ad}H_m^{int} = H_m^{\sigma} - H_m^{g} \tag{9.9}$$

and the **integral molar entropy of adsorption**

$$\Delta_{ad}S_m^{int} = S_m^{\sigma} - S_m^{g} \tag{9.10}$$

They are defined in complete analogy to the integral molar energy. The difference between the energy and the enthalpy of adsorption is usually small. If we treat the free gas as being ideal, the difference is $\Delta_{ad}U_m^{int} = \Delta_{ad}H_m^{int} + RT$. At 25°C $RT$ is only 2.4 kJ/mol. For this reason we do not need to worry too much about whether a heat of adsorption is the adsorption enthalpy or the internal adsorption energy, if we only want to estimate is.

Let us now consider how these quantities are related to experimentally determined heats of adsorption. An essential factor is the condition under which the calorimetric experiment is carried out. Under constant volume conditions, $\Delta_{ad}U_m^{int}$ is equal to the total heat of adsorption. In such an experiment a gas reservoir of constant volume is connected to a constant volume adsorbent reservoir (Fig. 9.3). Both are immersed in the same calorimetric cell. The total volume remains constant and there is no volume work. The heat exchanged equals the integral molar energy times the amount of gas adsorbed:

$$Q = \Delta_{ad}U_m^{int} \cdot N^{\sigma} \tag{9.11}$$

In general, $\Delta_{ad}U_m^{int}$ is negative (otherwise the substance would not adsorb) and heat is released upon adsorption.

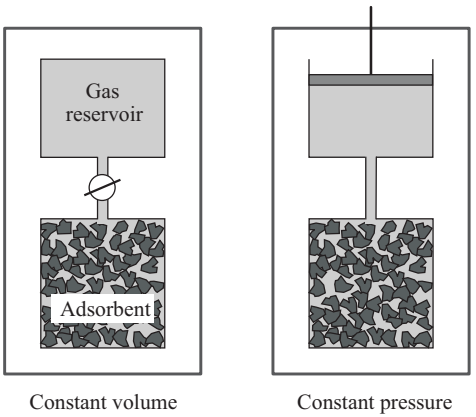



**Figure 9.3:** Schematic drawing of calorimeters for measuring heats of adsorption under constant volume and constant pressure conditions. The active volume is filled with the adsorbent usually in the form of a powder.

In practice, most calorimetric experiments are carried out under constant pressure. By moving the piston down, the pressure in the cell is increased and the heat release is measured. In this case the heat exchanged is equal to the integral enthalpy of adsorption :

$$Q = \Delta_{ad}H_m^{int} \cdot N^{\sigma} \tag{9.12}$$

The integral molar entropy of adsorption is obtained from a well-known thermodynamic relation: for a reversible, isothermal process the heat is equal to the change in entropy multiplied by the temperature. This directly leads to

$$\Delta_{ad}S_m^{int} = \frac{\Delta_{ad}H_m^{int}}{T} \tag{9.13}$$

One word about the Gibbs energies of adsorption. In equilibrium the molar Gibbs energy of adsorption is zero: $\Delta_{ad}G_m = \mu^{\sigma} - \mu^{g} = 0$. The reason is simple. In equilibrium and for constant $P$ and $T$ the chemical potential of the molecules in the gas phase $\mu^g$ is equal to the chemical potential of adsorbed molecules $\mu^{\sigma}$. What is not zero is the **standard Gibbs energy of adsorption**

$$\Delta_{ad}G_m^0 = \mu^{0\sigma} - \mu^{0g} \tag{9.14}$$

It is a molar quantity but since the name is so long the "molar" is not explicitly stated. The problem with $\Delta_{ad}G_m^0$ is that we cannot measure it directly but its value depends on the specific model of adsorption used. We return to it when introducing the Langmuir model.

### 9.2.2 Differential quantities of adsorption and experimental results

We have introduced integral molar quantities, which indicates that there are corresponding differential quantities. Integral refers to the fact that the total amount of adsorbed gas is involved. In contrast, the **differential molar energy of adsorption** is determined only by the last infinitesimal amount adsorbed. It is defined as

$$\Delta_{ad}U_m^{dif} = \left.\frac{dU^\sigma}{dN^\sigma}\right|_{T,A} - \left.\frac{dU^g}{dN^\sigma}\right|_{T,A} \tag{9.15}$$

$U^g$ is the total internal energy of the free gas. Since usually the amount adsorbed is small compared to the total amount of gas in the reservoir the properties of the free gas do not change significantly during adsorption. Thus, $dU^g/dN^\sigma = U_m^g$ which leads to

$$\Delta_{ad}U_m^{dif} = \left.\frac{dU^\sigma}{dN^\sigma}\right|_{T,A} - U_m^g \tag{9.16}$$

It involves the change of the internal surface energy upon adsorption of an infinitesimal amount of gas at constant temperature and total surface area.

We have to distinguish between integral and differential quantities because the energy changes with the amount adsorbed. This can have at least three causes: First, most surfaces are energetically heterogeneous and the binding sites with a high binding energy are occupied first. Second, in general the first monolayer has a different binding energy from the next layer because its adsorption is dominated by the interaction of the solid adsorbent with the gas molecule. For the second layer the interaction between adsorbed gas molecules with gas molecules is important. Third, if molecules interact laterally with neighboring molecules on the surface it is energetically more favorable for molecules to adsorb to a partially covered surface.

By analogy, the **differential molar enthalpy of adsorption** and the **differential molar entropy of adsorption** are defined as:

$$\Delta_{ad}H_m^{dif} = \left.\frac{dH^\sigma}{dN^\sigma}\right|_{T,\gamma} - H_m^g \tag{9.17}$$

$$\Delta_{ad}S_m^{dif} = \left.\frac{dS^\sigma}{dN^\sigma}\right|_{T,A} - S_m^g \tag{9.18}$$

Physical adsorption of gases on solids is virtually always enthalpically driven ($\Delta_{ad}H_m^{dif} < 0$). Entropically driven adsorption can exist but usually the entropy of molecules on a surface is much lower than in the gas phase. Vibrational, rotational, and also translational degrees of freedom are restricted on surfaces.

■ **Example 9.1.** As an example, a typical adsorption isotherm for benzene adsorbing to graphitized carbon blacks is shown in Fig. 9.4 [369]. Graphitized carbon blacks are produced by heating carbon in the absence of air to 3000°C. The initially spherical carbon particles become polyhedral with faces consisting mainly of homogeneous basal planes of single crystal graphite. In the adsorption isotherm, three regimes can be distinguished:

- At very low pressures ($P/P_0 < 0.1$) the adsorption isotherm rises steeply. Adsorbing molecules find many free binding sites. These very few molecules on the surface have a chance to bind to strong binding sites at grain boundaries. This can also be seen from the differential heats of adsorption: at coverages below 0.3 $\mu$mol/m$^2$ the heat of adsorption is maximal.

- Monolayer coverage is reached at a pressure of $P/P_0 \approx 0.1$. At this point the steep slope of the adsorption isotherm levels off. For the first monolayer, a roughly constant heat of adsorption of 43 kJ/mol is observed. This is about 9 kJ/mol higher than the heat of condensation of benzene.

- At higher vapor pressures $P/P_0 > 0.1$) multilayers are formed. In the multilayer region the slope becomes steeper again with increasing pressure. For $P \to P_0$ the adsorbed layer gets very thick because macroscopic condensation sets in. The differential heat of adsorption is slightly above the heat of condensation, but significantly lower than the value for the first monolayer.

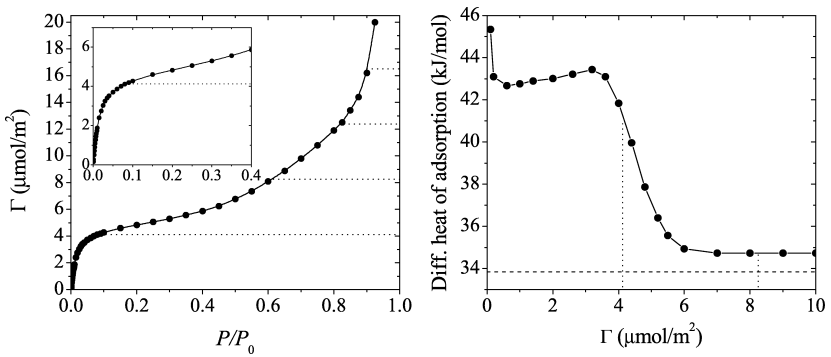



**Figure 9.4:** Left: Adsorption isotherm for benzene ($C_6H_6$) adsorbing to graphitized thermal blacks at 20°C. The insert shows the adsorption isotherm for low coverages in more detail. Dotted lines indicate mono- or multilayer coverages at multiples of 4.12 $\mu$mol/m$^2$. The equilibrium vapor pressure of benzene at 20°C is $P_0 = 10.2$ kPa. Right: Differential heat of adsorption versus adsorbed amount. The dashed line corresponds to the heat of condensation of bulk benzene. Redrawn after Ref. [369].

# 9.3 Adsorption models

## 9.3.1 The Langmuir adsorption isotherm

A simple model to describe adsorption was presented by Langmuir [370]. Langmuir assumed that on the surface there are a certain number of binding sites per unit area $S$ (fig. 9.5). $S$ is in units of mol/m$^2$ (or m$^{-2}$). Of these binding sites $S_1$ are occupied with adsorbate and $S_0 = S - S_1$ are vacant. The adsorption rate in moles per second and per unit area

is proportional to the number of vacant binding sites $S_0$ and to the pressure: $k_{ad}PS_0$. The desorption rate is proportional to the number of adsorbed molecules $S_1$ and equal to $k_{de}S_1$, where $k_{de}$ is a constant. In equilibrium the adsorption rate must be equal to the desorption rate; otherwise the number of adsorbed molecules would change. Therefore

$$k_{de}S_1 = k_{ad}PS_0 = k_{ad}P \cdot (S - S_1) \tag{9.19}$$

$$\Rightarrow k_{de}S_1 + k_{ad}PS_1 = k_{ad}PS \Leftrightarrow \frac{S_1}{S} = \frac{k_{ad}P}{k_{de} + k_{ad}P} \tag{9.20}$$

$S_1/S$ is the coverage $\theta$. With $K_L = k_{ad}/k_{de}$ we get the Langmuir equation :

$$\theta = \frac{K_L P}{1 + K_L P} \tag{9.21}$$

The Langmuir constant is given in Pa$^{-1}$, $k_{ad}$ is given in s$^{-1}$Pa$^{-1}$, and $k_{de}$ is in s$^{-1}$.

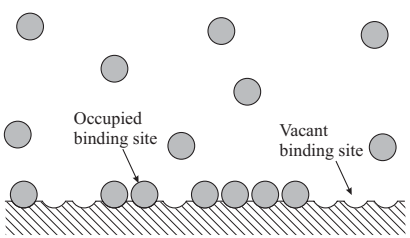



**Figure 9.5:** Schematic drawing of the Langmuir adsorption model.

Typical Langmuir adsorption isotherms are plotted in figure 9.6 for different values of the Langmuir constant. If adsorption from solution is considered, the pressure $P$ has to be replaced by the concentration $c$ and the Langmuir constant is given in units of L mol$^{-1}$ instead of Pa$^{-1}$.

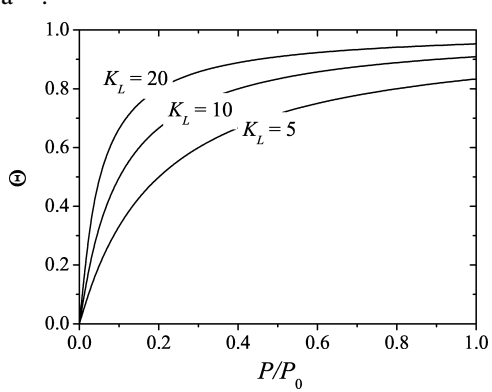



**Figure 9.6:** Langmuir adsorption isotherms plotted as coverage $\theta$ versus relative vapor pressure for three different Langmuir constants. The pressure and the Langmuir constants are normalized by dividing them by $P_0$.

Alternatively, the Langmuir adsorption isotherm equation can be expressed by the number of adsorbed moles per gram or surface area

$$\Gamma = \frac{\Gamma_{mon} K_L P}{1 + K_L P} \tag{9.22}$$

Here, $\Gamma_{mon}$ is the number of adsorbed moles per gram or per unit area of substrate, when all binding sites are occupied and a monolayer of molecules is bound. $\Gamma_{mon}$ is related to the surface area occupied by one adsorbed molecule $\sigma_A$ by $\Gamma_{mon} = \Sigma/\left(N_0\sigma_A\right)$ or $\Gamma_{mon} = 1/(N_0\sigma_A)$.

What is the significance of the constants $k_{ad}$ and $k_{de}$? $k_{de}$ is the inverse of the adsorption time:

$$k_{de} = \frac{1}{\tau_0} \cdot e^{-Q/k_BT} \tag{9.23}$$

In order to calculate $k_{ad}$ we remember the kinetic theory of ideal gases and Eq. (2.1) of example 2.1. Equation (2.1) tells us how many gas molecules of mass $m$ hit a certain area $A$ per second. If we take the area to be the active area of one binding site $\sigma_A$ the number of molecules hitting one binding site per second is

$$\frac{\sigma_A P}{\sqrt{2\pi m k_B T}} \tag{9.24}$$

If we assume that each molecule that hits the surface sticks to it, expression (9.24) is equal to $k_{ad}P$ and we get

$$k_{ad} = \frac{\sigma_A}{\sqrt{2\pi m k_B T}} \tag{9.25}$$

For the Langmuir constant we obtain

$$K_L = K_L^0 \cdot e^{Q/k_BT} \qquad \text{with} \qquad K_L^0 = \frac{\sigma_A \tau_0}{\sqrt{2\pi m k_B T}} \tag{9.26}$$

■ **Example 9.2.** Let us estimate the Langmuir constant for the physisorption of a gas to a solid surface at a temperature of 120 K. We take $\tau_0$ to be $10^{-13}$ s, use a typical molecular cross-section of 10 Å$^2$, and assume a heat of adsorption of 20 kJ/mol. As a gas we consider nitrogen ($M = 0.028$ kg/mol). It is more convenient to use Eq. (9.26) in molar rather than molecular units:

$$K_L = K_L^0 \cdot e^{Q/RT} \qquad \text{with} \qquad K_L^0 = \frac{N_A \sigma_A \tau_0}{\sqrt{2\pi M R T}} \tag{9.27}$$

Inserting the values leads to $K_L^0 = 4.55 \times 10^{-10}$ Pa$^{-1}$ and $K_L = 0.23$ Pa$^{-1}$.

Please keep in mind, that the calculated adsorption rate is an upper limit. It can also be used for the calculation of the condensation rate of a liquid. We define a **condensation coefficient** or **sticking probability** as the ratio between the actual condensation and the (just calculated) upper limit. The sticking probability can be determined from molecular beam experiments. Examples: For $N_2$ on tungsten the sticking probability at 27°C is 0.61 [372]. For $H_2O$ on ice at 200 K it is close to one [371]. The sticking probability can also be significantly below 1. Then the Langmuir constant is reduced accordingly.

The kinetic derivation has the disadvantage that it refers to a certain model. The Langmuir adsorption isotherm, however, applies under more general conditions and it is possible to derive it with the help of statistical thermodynamics [8,373]. Necessary and sufficient conditions for the validity of the Langmuir equation (9.21) are:

- The molecules bind to well-determined binding sites on the adsorbent.
- Each binding site can bind only one molecule.
- The binding energy is independent of the presence of other bound molecules.

### 9.3.2 The Langmuir constant and the Gibbs energy of adsorption

The Gibbs energy of adsorption depends on the specific model used (Section 9.2.1). Here, we demonstrate this with the Langmuir model and derive a relation between the standard Gibbs energy of adsorption $\Delta_{ad}G_m^0$ and the Langmuir constant. Therefore, we treat the binding and desorption of gas molecules to surface binding sites, like a chemical reaction. Chemical equilibria are commonly characterized by an equilibrium constant $K$. For the dissociation reaction $\mathrm{AB} \rightleftharpoons \mathrm{A} + \mathrm{B}$ this constant is given by

$$K = \frac{[A] \cdot [B]}{[AB]} \tag{9.28}$$

It is related to the standard Gibbs energy of the reaction $\Delta_r G_m^0$ by

$$\Delta_r G_m^0 = -RT \cdot \ln K \tag{9.29}$$

Here, $[AB]$, $[A]$, and $[B]$ are the concentrations (for gas reactions, the pressures) of the bound (educt) and dissociated molecules (products), respectively.

Let us apply this formalism to the adsorption of gas molecules to a surface. The adsorption equilibrium constant in units of Pa is

$$K_{ad} = \frac{S_0 P}{S_1} \tag{9.30}$$

It is related to the standard Gibbs energy of adsorption by

$$\Delta_{ad} G_m^0 = -RT \cdot \ln K_{ad} \tag{9.31}$$

Inserting $S - S_1 = S_0$ leads to

$$K_{ad} = \frac{(S - S_1)\, P}{S_1} \Rightarrow S_1 = \frac{SP}{K_{ad} + P} \tag{9.32}$$

Since $\theta = S_1/S$ we immediately get

$$\theta = \frac{P}{K_{ad} + P} \qquad \text{with} \qquad K_{ad} = \exp\left(-\frac{\Delta_{ad} G_m^0}{RT}\right) \tag{9.33}$$

Comparison with Eq. (9.21) shows that

$$K_L = \frac{1}{K_{ad}} \tag{9.34}$$

When inserting real numbers in Eqs. (9.31) and (9.33) it is important to remember that pressures are given in units of normal pressure, that is $10^5$ Pa.

### 9.3.3 Langmuir adsorption with lateral interactions

One assumption of the Langmuir model is that the adsorbed molecules do not interact with each other. It is possible to modify the theory and take such an interaction into account. Therefore we assume that each binding site has $n$ neighboring binding sites. The average number of neighbors of an adsorbed molecule is $n\theta$. When we denote the additional binding energy related to the interaction between a pair of neighboring molecules by $E_P$, we can consider lateral interactions by modifying the Langmuir equation. Therefore it is convenient to write the Langmuir equation as

$$\frac{\theta}{1-\theta} = K'_L P$$

and modify the Langmuir constant:

$$K'_L = K^0_L \cdot \exp\left(\frac{Q + nE_P\theta}{RT}\right) = K_L \cdot \exp\left(\frac{nE_P\theta}{RT}\right) \tag{9.35}$$

This equation is sometimes called the Frumkin–Fowler–Guggenheim (FFG) isotherm [374–376]. For $\beta \equiv nE_P/RT < 4$ lateral interactions cause a steeper increase of the adsorption isotherm in the intermediate pressure range. Characteristic of all Langmuir isotherms is a saturation at high partial pressures $P/P_0 \to 1$.

A remarkable shape is calculated with Eq. (9.35) for $\beta > 4$. A region is obtained where the $\theta$-versus-$P$ curve has a negative slope (dotted curve in Fig. 9.7). This is physically nonsense: The coverage is supposed to decrease with increasing pressure and for one pressure there are three possible values of $\theta$. In reality this is a region of two-phase equilibrium. Single adsorbed molecules and clusters of adsorbed molecules coexist on the surface. The situation is reminiscent of the three-dimensional van der Waals equation of state which can be used to describe condensation.

As an example, Fig. 9.7 shows adsorption isotherms of krypton on the (0001) face of graphite. The dashed lines were fitted using Eq. (9.35) with $\beta = 4.5$. In reality the coverage increases steeply and the two-phase region can be identified. Figure 9.7 shows another typical feature of adsorption: The amount adsorbed decreases with increasing temperature.

### 9.3.4 The BET adsorption isotherm

In Langmuir model, the maximal adsorption is that of a monolayer. Langmuir adsorption isotherms all saturate at high vapor pressures. This is unrealistic for many cases. In order to consider the adsorption of multilayers, Brunauer, Emmett, and Teller extended the Langmuir theory and derived the so-called BET adsorption isotherm [378]. The basic idea in the BET theory was to assume a Langmuir adsorption for each of the layers (Fig. 9.8).

It is assumed that the adsorption heat for the first layer $Q_1$ has a particular value. For all further layers, the heat of adsorption $Q_i$ corresponds to the heat of condensation of the liquid. Another condition is that desorption and adsorption take place only directly between vapor and surface. Adsorbed molecules are not allowed to move from one layer directly to another. In equilibrium, the desorption rate for each layer must be equal to the adsorption rate. We

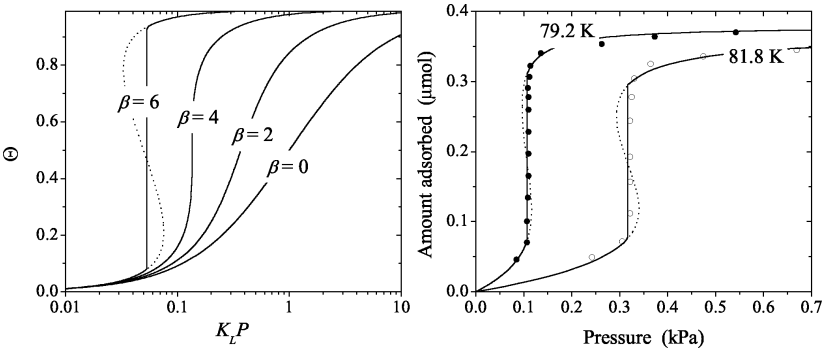



**Figure 9.7:** Left: Frumkin–Fowler–Guggenheim (FFG) adsorption isotherms (coverage $\theta$ versus the pressure in units of $K_L^{-1}$). The curves were calculated using Eq. (9.35) with $\beta = 0, 2, 4, 6$. For $\beta = 6$ the physically correct adsorption curve is plotted as a continuous curve while the one calculated with Eq. (9.35) is plotted as a dotted curve. Right: Adsorption isotherms for krypton adsorbing to the (0001) plane of graphite at two different temperatures. The dotted curves were fitted using Eq. (9.35) with $\beta = 4.5$. Experimental results were taken from Ref. [377].

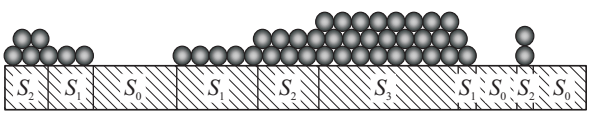



**Figure 9.8:** BET model of adsorption.

have the following adsorption and desorption rates:

$$\begin{array}{lll} \text{Adsorption to vacant surface sites} & k_{ad}^1 P S_0 & \\ \text{Desorption from first layer} & & = a_1 S_1 e^{-Q_1/RT} \\ \text{Adsorption to the } i^{th} \text{ layer} & k_{ad}^i P S_{i-1} & \\ \text{Desorption from the } i^{th} \text{ layer} & & = a_i S_i e^{-Q_i/RT} \end{array} \tag{9.36}$$

Here, $a_1$ and $a_i$ are frequency factors like $1/\tau_0$. As a result we get

$$\frac{n}{n_{mon}} = \frac{C}{\left(1 - \frac{P}{P_0}\right) \cdot \left[1 + \frac{P}{P_0} \cdot (C-1)\right]} \cdot \frac{P}{P_0} \tag{9.37}$$

Here, $n$ is the total number of moles adsorbed per unit area, $n_{mon}$ is the number of adsorbed moles in one full monolayer per unit area (each binding site is occupied exactly once), $P_0$ is the equilibrium vapor pressure, and

$$C = \frac{a_i k_{ad}^1}{a_1 k_{ad}^i} \cdot e^{(Q_1 - Q_i)/RT} \approx e^{(Q_1 - Q_i)/RT} \tag{9.38}$$

Equation (9.37) shows that, $n/n_{mon}$ becomes infinite for $P/P_0 \to 1$. This is what we expect because condensation sets in. For the second part of Eq. (9.38) see exercises.

Since the BET adsorption isotherm is so widely used, we describe a simple derivation [1]. It is convenient to define two parameters, $\alpha$ and $\beta$, according to

$$\alpha = \frac{k_{ad}^1 P}{a_1}\, e^{Q_1/RT} \qquad \text{and} \qquad \beta = \frac{k_{ad}^i P}{a_i}\, e^{Q_i/RT} \tag{9.39}$$

Using these parameters we can write

$$S_1 = \alpha S_0 \qquad \text{and} \qquad S_2 = \beta S_1 \tag{9.40}$$

and in general

$$S_i = \beta^{i-1} S_1 = \alpha \beta^{i-1} S_0 = C\beta^i S_0 \tag{9.41}$$

The number of moles adsorbed per unit area is given by

$$n = S_1 + 2S_2 + 3S_3 + \ldots = \sum_{i=1}^{\infty} i S_i = C S_0 \sum_{i=1}^{\infty} i\beta^i = \frac{C S_0 \beta}{(1-\beta)^2} \tag{9.42}$$

Monolayer coverage can be written as

$$n_{mon} = S_0 + S_1 + S_2 + S_3 + \ldots = S_0 + C S_0 \sum_{i=1}^{\infty} \beta^i = S_0 + \frac{C S_0 \beta}{1-\beta} \tag{9.43}$$

Now we divide Eq. (9.42) by Eq. (9.43)

$$\frac{n}{n_{mon}} = \frac{C S_0 \beta / (1-\beta)^2}{S_0 + C S_0 \beta / (1-\beta)} = \frac{C\beta}{(1-\beta)\left[1-\beta+C\beta\right]} \tag{9.44}$$

An essential step is to realize that the factor $\beta$ is equal to $P/P_0$. To see this let us consider the adsorption of a vapor to its own liquid at equilibrium and hence at $P_0$. We take this situation to be similar to the situation of the vapor adsorbing to the $i$th (not the first) layer. Taking the rate of binding equal to the rate of desorption we write

$$k_{ad}^i P_0 S = a_i S e^{-Q_i/RT} \tag{9.45}$$

On both sides we insert the total number of binding sites per unit area $S$ because all sites can serve as binding and desorption sites. This immediately leads to

$$P_0 = \frac{a_i}{k_{ad}^i}\, e^{-Q_i/RT} \tag{9.46}$$

By inserting this into Eq. (9.39) we get $\beta = P/P_0$ and with Eq. (9.44) obtain the BET isotherm equation (9.37).

Figure 9.9 shows how BET isotherms depend on the parameter $C$. For high values of $C$ the binding of vapor molecules directly to the surface, is strong compared to the intermolecular interaction. Therefore, at least for low pressures, a Langmuir type of adsorption is obtained. Only at high pressures do the molecules start to form multilayers. For low values of $C$ the

molecules prefer binding to themselves while the binding energy to the surface is low. Therefore, the first monolayer only forms at relatively high pressures. Once it has formed it is easier for the next molecules to adsorb.

BET isotherms are widely used to fit experimental adsorption isotherms. Two examples, the adsorption of water to alumina and silica, are shown in Fig. 9.9 on the right. The adsorption of alumina can be fitted with the BET equation up to a relative pressure of 0.4, which is quite typical. For silica the fit is acceptable even up to $P/P_0 = 0.8$.

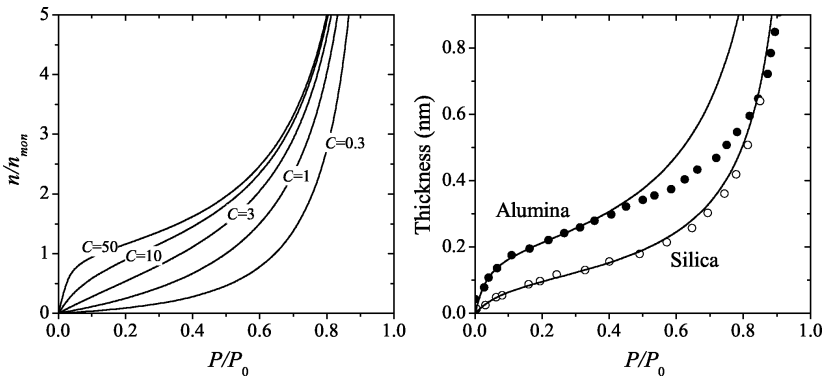



**Figure 9.9:** Left: BET adsorption isotherms plotted as total number of moles adsorbed, $n$, divided by the number of moles in a complete monolayer, $n_{mon}$, versus the partial pressure, $P$, divided by the equilibrium vapor pressure, $P_0$. Isotherms were calculated for different values of the parameter $C$. Right: Adsorption isotherms of water on a sample of alumina (Baikowski CR 1) and silica (Aerosil 200) at 20°C ($P_0 = 2.7$ kPa, redrawn from Ref. [379]). The BET curves were plotted using Eq. (9.37) with $C = 28$ (alumina) and $C = 11$ (silica). To convert from $n/n_{mon}$ to thickness, the factors 0.194 nm and 0.104 nm were used, which correspond to $n_{mon} = 6.5$ and 3.6 water molecules per nm$^2$, respectively.

### 9.3.5 Adsorption on heterogeneous surfaces

Surfaces are usually not perfectly homogeneous. Different crystal faces are exposed, defects and other deviations from the perfect lattice are present. Often there are different types of molecules as in steel (e.g. Fe, C, Ni, Co) or in glass (e.g. $SiO_2$, B, Na, K) and their concentrations on the surface might vary locally.

On heterogeneous surfaces the binding energy of an adsorbate will generally not be a fixed value, but there is a distribution of binding energies. The probability of finding a binding site in the energy interval $Q \ldots Q + dQ$ is described by a distribution function $f(Q)dQ$. The experimentally observed adsorption is the sum of all adsorption events on all different kinds of binding sites. At a fixed temperature the coverage is

$$\theta(P) = \int_0^\infty \theta^H(Q, P) \cdot f(Q) \cdot dQ \tag{9.47}$$

The distribution function is normalized:

$$\int_0^\infty f(Q)dQ = 1 \tag{9.48}$$

For the adsorption isotherm on a well determined homogeneous part of the surface $\theta^H(Q, P)$ the Langmuir equation is often used.

A relatively well-known adsorption isotherm, the Freundlich isotherm [366]

$$\theta = \left(\frac{P}{P_0}\right)^{\frac{k_B T}{2Q^*}}, \tag{9.49}$$

is obtained with an exponentially decaying distribution of adsorption sites according to $f(Q) \propto e^{-Q/Q^*}$ and assuming a Langmuir behavior for $\theta^H$ [380]. Here, $Q^*$ is a constant which characterizes the distribution of adsorption energies. A requirement in the derivation is that $Q^* > k_B T$. This, however, is not a severe restriction because for $Q^* < k_B T$ adsorption is negligible.

Often one would like to conclude something about the distribution of the binding energies using a measured adsorption isotherm. This is difficult. Usually certain assumptions concerning $\theta^H$ have to be made [381].

### 9.3.6 The potential theory of Polanyi

Polanyi[3] approached the phenomenon of adsorption in a completely different way. He assumes that the molecules close to the surface feel a potential — similar to the gravitation field of the earth. One cause for this potential is the van der Waals attraction. The potential compresses the gas close to the surface, isothermally. Once the pressure becomes higher than the equilibrium vapor pressure, it condenses. In equilibrium, the chemical potential of the gas at a distance $x$ from the surface $\mu(x, P_x)$ must be equal to the chemical potential $\mu(\infty, P)$ at a large distance away form the surface. $P_x$ is the local pressure at a distance $x$, $P$ is the partial pressure far from the surface. The differential of the chemical potential is

$$d\mu = -S_m dT + V_m dP_x + dU_m \tag{9.50}$$

$S_m$ is the molar entropy, $V_m$ the molar volume of the gas at a distance $x$, and $U_m$ is the molar internal energy. Integrating the left side from infinite distance to a distance $x$ leads to

$$\int_{\mu(\infty,P)}^{\mu(x,P_x)} d\mu = \mu(x, P_x) - \mu(\infty, P) = 0 \tag{9.51}$$

The integration is simple since the chemical potential is a state function. The integral is zero because in equilibrium the chemical potential should be the same everywhere. Integration of the right side of Eq. (9.50) at constant temperature leads to

$$\int_P^{P_x} V_m dP' + U_m(x) - U_m(\infty) = 0 \Rightarrow -U_m(x) = \int_P^{P_x} V_m dP' \tag{9.52}$$

[3] Michael Polanyi, 1891–1976. Hungarian physicist who worked in Berlin and Manchester. His son, John Charles Polanyi, received the Noble Price in chemistry in 1986.

The internal energy at $x = \infty$ was set to zero: $U_m(\infty) = 0$. Now we assume that the gas behaves like an ideal gas until it condenses. For an ideal gas we can use $V_m P' = RT$ and solve the integral:

$$-U_m(x) = RT \cdot \ln \frac{P_x}{P}. \tag{9.53}$$

To take a possible condensation of the gas into account we assume that the gas becomes liquid once the pressure exceeds the equilibrium vapor pressure $P_0$. The amount adsorbed is just the thickness of the liquid film multiplied by its density. The film thickness $x_f$ is related to the number of adsorbed moles per unit area by $\Gamma = x_f/V_m^L$, where $V_m^L$ is the molar volume of the liquid. The energy at which the pressure $P_0$ is reached is

$$U_m(x_f) = -RT \cdot \ln \frac{P_0}{P} \tag{9.54}$$

As an example we consider an ideal gas, which is attracted towards a surface by van der Waals interactions. In Eq. (6.13) we calculated the Gibbs free energy for the van der Waals interaction of a single molecule with a flat surface. The internal energy is a factor of two higher. Using $C \equiv \pi \rho_B C_{AB}/3$ as a constant we obtain

$$-U_m(x_f) = \frac{C}{(D_0 + x_f)^3} = RT \cdot \ln \frac{P_0}{P} \Rightarrow x_f = \sqrt[3]{\frac{C}{RT \cdot \ln (P_0/P)}} - D_0 \tag{9.55}$$

$D_0$ is a distance, which corresponds to the molecular radius. The number of moles adsorbed per unit area is

$$\Gamma = \frac{x_f}{V_m^L} = \frac{1}{V_m^L} \cdot \sqrt[3]{\frac{C}{RT \cdot \ln (P_0/P)}} - \frac{D_0}{V_m^L} \tag{9.56}$$

■ **Example 9.3.** Plot the estimated adsorption isotherm for water vapor on silicon oxide at 20°C. First we need to estimate the constant $C$. From Chapter 6 we know that it is related to the Hamaker constant $A_H$ (Eq. 6.16): $C \equiv \pi \rho_B C_{AB}/3 = A_H/3\pi\rho_A$. Here, $\rho_A$ and $\rho_B$ are the number densities of molecules in liquid water and silicon oxide, respectively. The Hamaker constant for water interacting with silicon oxide across air is $A_H = 10^{-20}$ J (Table 6.3). With a density of water of 1000 kg/m$^3$, a molecular weight of 18 g/mol, and a molecular radius of $D_0 \approx 1$ Å we get $V_m^L = 0.018\,\mathrm{kgmol}^{-1}/(1000\,\mathrm{kgm}^{-3}) = 18 \times 10^{-6}\,\mathrm{m}^3\mathrm{mol}^{-1}$ and

$$C = \frac{A_H}{3\pi\rho_A} = \frac{10^{-20}\,\mathrm{J}}{3\pi \cdot 1000\,\mathrm{kgm}^{-3}/0.018\,\mathrm{kgmol}^{-1}} = 1.9 \times 10^{-26}\,\frac{\mathrm{Jm}^3}{\mathrm{mol}}$$

$$\begin{aligned} \Gamma &= \frac{1}{V_m^L} \cdot \sqrt[3]{\frac{C}{RT \cdot \ln (P_0/P)}} - \frac{D_0}{V_m^L} \\ &= 1.1 \times 10^{-5}\,\frac{\mathrm{mol}}{\mathrm{m}^2} \cdot \sqrt[3]{\frac{1}{\ln (P_0/P)}} - 0.55 \times 10^{-5}\,\frac{\mathrm{mol}}{\mathrm{m}^2} \end{aligned} \tag{9.57}$$

Figure 9.10 shows the calculated adsorption isotherm and compares it to experimental results. The Polanyi theory predicts the general shape of the adsorption isotherm reasonably well considering that it is not a fit but a real calculation. Real adsorption is, however, stronger, which indicates that in addition to van der Waals forces other attractive forces are important.

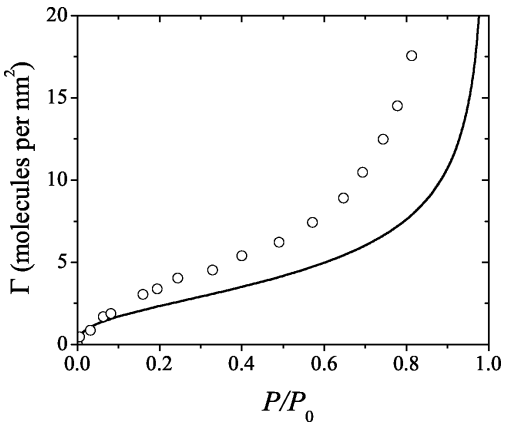



**Figure 9.10:** Adsorbed amount of water on a silicon oxide surfaces versus relative vapor pressure at 20°C. The continuous line was calculated with the theory of Polanyi and assuming van der Waals forces only (Eq. 9.57). Experimental results as measured on Aerosil 200 were adapted from Ref. [379] (see also Fig. 9.9). The deviation at high pressure is partially due to the porosity of the adsorbent. The equilibrium vapor pressure is $P_0 = 3.17$ kPa.

Which theory is suitable for a certain application? The adsorption theory of Henry is applicable at low pressure. This, however, is natural since it can be viewed as the first term in a series of the adsorption function. A widely used adsorption isotherm equation is the BET equation. It usually fits experimental results for $0.05 < P/P_0 < 0.35$. For very small pressures the fit is not perfect due to the heterogeneity. For higher pressures the potential theory is more suitable at least for flat, homogeneous adsorbents. It often applies to $P/P_0$ values from 0.1 to 0.8. Practically for $P/P_0 > 0.35$ adsorption is often dominated by the porosity of the material. A more detailed description of adsorption is obtained by computer simulations [382].

# 9.4 Experimental aspects of adsorption from the gas phase

## 9.4.1 Measurement of adsorption isotherms

Several experimental methods are applied to measure adsorption isotherms . The main problem is to determine the amount adsorbed. One method is a **gravimetric measurement**. In a gravimetric measurement the weight increase as a function of the pressure is determined. The adsorbent, usually in the form of a powder, is placed into a bulb and kept at the desired temperature. The bulb is mounted on a sensitive balance. Before the experiment, the bulb is

evacuated. Then the gas of interest is admitted into the bulb at a certain pressure. The increase in weight is measured. Dividing by the total surface area of the adsorbent, we get the amount of adsorbed gas. The pressure is increased and the weight measurement is repeated. In this way a whole isotherm is recorded.

In **volumetric measurements** the volume of an adsorbed gas at constant pressure and temperature is determined. Therefore, we first determine the "dead space" or volume of the bulb by admitting some nonadsorbing (or weakly adsorbing) gas such as Helium. Then, after evacuating the bulb, the gas of interest is admitted into the bulb. This is done at constant pressure and temperature. The volume admitted into the bulb minus the dead space is the amount adsorbed.

More recently methods have also been developed to measure the adsorbed amount on single surfaces and not onto powders. Adsorption to isolated surfaces can, for instance, be measured with a **quartz crystal microbalance** (QCM) [383]. The quartz crystal microbalance consists of a thin quartz crystal that is plated with electrodes on the top and bottom (Fig. 9.11). Since quartz is a piezoelectric material, the crystal can be deformed by an external voltage. By applying an AC voltage across the electrodes, the crystal can be excited to oscillate in a transverse shear mode at its resonance frequency. This resonance frequency is highly sensitive to the total oscillating mass. For an adsorption measurement, the surface is mounted on such a quartz crystal microbalance. Upon adsorption, the mass increases, which lowers the resonance frequency. This reduction of the resonance frequency is measured and the mass increase is calculated [384–387].

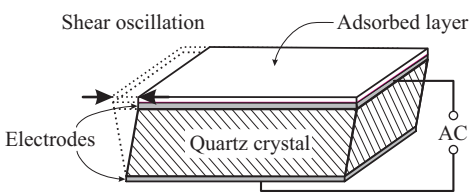



**Figure 9.11:** Working principle of a quartz crystal microbalance.

Another technique applicable to single flat surfaces is **ellipsometry**. To understand ellipsometry we first have to remember that the reflection of light from a surface depends on the direction of polarization. Light polarized parallel and perpendicular to the surface is reflected differently. Amplitude and phase change in a different way depending on polarization. In ellipsometry we direct a monochromatic and polarized light beam, usually that of a laser, onto a surface at a certain angle. The angle of incidence must be small enough so that the light is totally reflected. At the interface the polarization changes. Two parameters, sometimes called the ellipsometric angles $\Delta$ and $\Psi$, are measured. The first parameter is $\Delta = \delta_{in} - \delta_{out}$, where $\delta_{in}$ is the phase difference between the parallel and perpendicular component of the incoming wave and $\delta_{out}$ is the phase difference between the parallel and perpendicular component of the outgoing wave. The second parameter is given by $\tan\Psi = |R_p| / |R_s|$, where $R_p$ and $R_s$ are the reflection coefficients for the parallel and perpendicular polarized light. From the amplitude ratios and the phase shift between the incident and reflected light, we can determine the thickness of a thin film, assuming that the refractive index of the material is known.

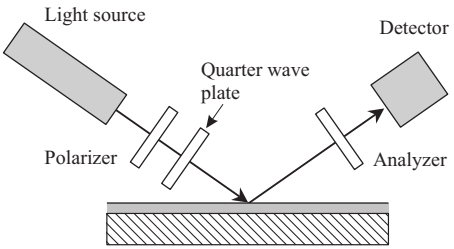



**Figure 9.12:** One type of ellipsometer, called a Null instrument. Polarizer and analyzer are rotated alternately until the null is found.

Different experimental configurations are used to determine the $\Delta$ and $\Psi$. A practical introduction into ellipsometry is Ref. [388]. One common type is the Null ellipsometer also called PCSA configuration for "polarizer–compensator–sample–analyzert't' (Fig. 9.12). In a Null ellipsometer elliptically polarized light produced by a laser, polarizer, and a quarter wave plate (or a compensator) is directed onto the surface. The polarization is adjusted by the polarizer in such a way, that perfectly linear polarized light is reflected from the sample. For linear polarized light the intensity can be decreased to zero by the analyzer. From the orientations of polarizer and analyzer at zero intensity (for given fixed properties of the compensator and angle of incidence) the ellipsometric angles can be calculated. As one example the adsotprion isotherm for nitrogen to the basal plane of graphite is shown in Fig. 9.13.

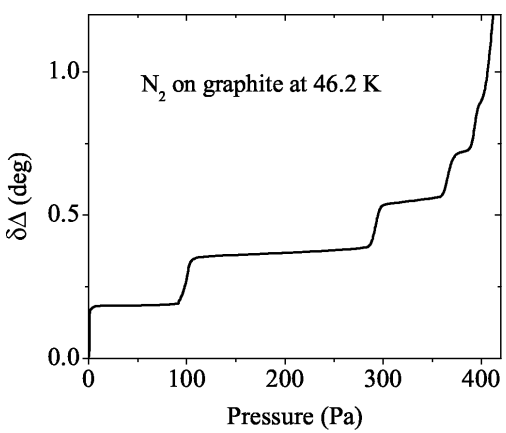



**Figure 9.13:** Adsorption of nitrogen to a single basal plane of graphite at a temperature of 46.2 K as determined by ellipsometry. Plotted is the change in the ellipsometric angle $\triangle$ versus pressure. The subsequent adsorption of at least four layers at defined pressures can clearly be distinguished. Redrawn from Ref. [389].

### 9.4.2 Procedures to measure the specific surface area

For all experiments with porous materials or powders we need to know the specific surface area. This is also important for many industrial applications and determination of the specific surface area has become a standard technique in chemical engineering of powders.

Usually specific surface areas are determined from adsorption experiments. To illustrate this let us assume that adsorption of a specific sample is adequately described by the Langmuir Eq. (9.22). From fitting experimental results we obtain $\Gamma_{mon}$ in units of mol/g. Then we assume a reasonable value for the cross-section area of a gas molecule $\sigma_A$, and obtain the specific surface from $\sum = \Gamma_{mon}\sigma_A N_A$. In most practical applications the BET adsorption isotherm is used instead of the Langmuir Eq. (9.22) because it fits better. From a fit with the BET isotherm we get $\Gamma_{mon}$ or $n_{mon}$. Some cross-sectional areas for suitable gases in Å$^2$ are: $N_2$: 16.2; $O_2$: 14.1; Ar: 13.8; $n$-$C_4H_{10}$: 18.1.

**Table 9.2:** Calculated specific surface areas $\Sigma$ in m$^2$/g for a batch of anatas, porous glass, a silica gel, and a special sample of the protein albumin [390].

| | $N_2$ | Ar | $O_2$ | CO | $CH_4$ | $CO_2$ | $C_2H_2$ | $NH_3$ |
|---|---|---|---|---|---|---|---|---|
| Anatas, $TiO_2$ | 13.8 | 11.6 | | 14.3 | | 9.6 | | |
| Porous glass | 232 | 217 | | | | 164 | 159 | 207 |
| Silica gel | 560 | 477 | 464 | 550 | | 455 | | |
| Albumin | 11.9 | 10.5 | 9.9 | | 10.3 | | | |

How do we practically determine the specific surface area? We measure the adsorption isotherm of a defined mass of adsorbent and fit it with the BET equation. Usually the BET model describes adsorption for $0.05 < P/P_0 < 0.35$ reasonably well, and we can restrict the measurement to that pressure range. Most commonly the volumetric technique is applied. As a result of our measurements, we get the volume of gas adsorbed, $V^{ad}$. To be meaningful the conditions (pressure, temperature) at which the volume of the gas is reported, have to be given. Usually standard conditions are chosen although the actual experiment is often done at lower temperature. For the analysis we have to transform the BET adsorption equation. First, we express the number of moles adsorbed by the volumes: $n/n_{mon} = V^{ad}/V^{ad}_{mon}$, where $V^{ad}_{mon}$ is the volume of gas required to get one complete monolayer. Inserting into Eq. (9.37) and rearrangement of leads to

$$\frac{P/P_0}{V^{ad}\left(1-P/P_0\right)} = \frac{1}{CV^{ad}_{mon}} + \frac{P/P_0 \cdot (C-1)}{CV^{ad}_{mon}} \tag{9.58}$$

As a result we plot $\frac{P/P_0}{V^{ad}(1-P/P_0)}$ versus $P/P_0$. This should give a straight line with a slope $(C-1)\,/\left(CV^{ad}_{mon}\right)$ and an intersection with the ordinate at $1/\left(CV^{ad}_{mon}\right)$. From the slope and intersection $C$ and $V^{ad}_{mon}$ can be determined.

Porous materials are often analyzed with a **mercury porosimeter**. With a mercury porosimeter we can measure the pore distribution of a solid. Thus, we can determine the specific surface area. Mercury is used because of its high surface tension (0.48 N/m) it does not wet

solid surfaces; typical contact angles are $140°$ on glass and $149°$ on paraffin. In the mercury porosimeter mercury is pressed with a pressure $P$ into the pores of a solid. Only pores whose radius is larger than

$$r > \frac{2\gamma_{Hg} \cdot |\cos\Theta|}{P} \tag{9.59}$$

are filled. We measure the filled volume versus the applied pressure. To get the wetted surface area $A$ we take the volume work and equate it with the work necessary to wet the surface:

$$PdV = -\left(\gamma_S - \gamma_{SL}\right) \cdot dA = -\gamma_{Hg} \cdot \cos\Theta \cdot dA \tag{9.60}$$

The pore is modelled as a capillary. Please note: $\gamma_{SL}$ is larger in this case than $\gamma_S$ and $\cos\Theta < 0$. Integration results in the entire surface area being wetted by mercury:

$$A = -\frac{1}{\gamma_{Hg} \cdot \cos\Theta} \cdot \int_0^{P_{\max}} PdV \tag{9.61}$$

The method is limited to pore radii larger than $\approx 2$ nm due to the maximal applicable pressure. An assumption made in the analysis is that a pore is accessible only by capillaries with larger radius. Pores with a narrow entrance and a wide body (called "ink bottle pores") lead to a hysteresis in the volume–pressure curve.

### 9.4.3 Adsorption on porous solids — hysteresis

Until now we have mainly treated adsorption onto non-porous surfaces. In reality, most industrial and many natural materials are porous: Textiles, paper, bricks, sand, porous rocks, food products, zeolites etc. We start our discussion with a classification of pores according to their size, which is recommended by IUPAC:

- Macropores have a diameter larger than 50 nm. Macropores are so wide that gases adsorb virtually to flat surfaces.
- Mesopores are in the range of 2–50 nm. Capillary condensation often dominates the filling of mesopores. Below the critical temperature, multilayers arise. Pores, on one hand, limit the number of layers, but on the other hand, capillary condensation can occur.
- Micropores are smaller than 2 nm. In micropores the structure of the adsorbed fluid is significantly different from its macroscopic bulk structure. Confined liquids are a highly active area of research because of their unique properties. An important example of a microporous materials are zeolites, which are used for catalysis.

This classification is certainly not perfect because the filling of pores is also determined by their shape (cylinders, slits, cones, irregular) and pores may be separate or connected.

To characterize the size of pores, the hydraulic radius $a_h$ was introduced. The hydraulic radius is the ratio of void volume and void area. For a long cylindrical pore of length $l$ and radius $r$ ($r \ll l$) the hydraulic radius is for instance

$$a_h = \frac{\pi r^2 l}{2\pi r l} = \frac{r}{2} \tag{9.62}$$

Above a certain critical temperature no multilayers should adsorb. Hence, even porous solids should behave like flat surfaces, if no pores of molecular size are present. Capillary condensation plays a role only below the critical temperature.

Adsorption to porous materials is often characterized by hysteresis in the adsorption behavior. Such a hysteresis is observed when, after the adsorption process, a desorption experiment is done in which the pressure is progressively reduced from its maximum value and the desorption isotherm is measured. During the desorption process, the liquid phase vaporizes from the pores. The desorption isotherm does not precisely track the adsorption isotherm, but lies above it. Moreover isotherms often flatten out with high $P/P_0$ values, because filling up of pores decreases the available surface area.

It is widely accepted that two mechanisms contribute to the observed hysteresis. The first mechanism is thermodynamic in origin [391,392]. It is illustrated in Fig. 9.14 for a cylindrical pore of radius $r_c$. The adsorption cycle starts at a low pressure. A thin layer of vapor condenses onto the walls of the pore (1). With increasing pressure the thickness of the layer increases. This leads to a reduced radius of curvature for the liquid cylinder $a$. Once a critical radius $a_c$ is reached (2), capillary condensation sets in and the whole pore fills with liquid (3). When decreasing the pressure again, at some point the liquid evaporates. This point corresponds to a radius $a_m$ which is larger than $a_c$. Accordingly, the pressure is lower. For a detailed discussion see Ref. [393].

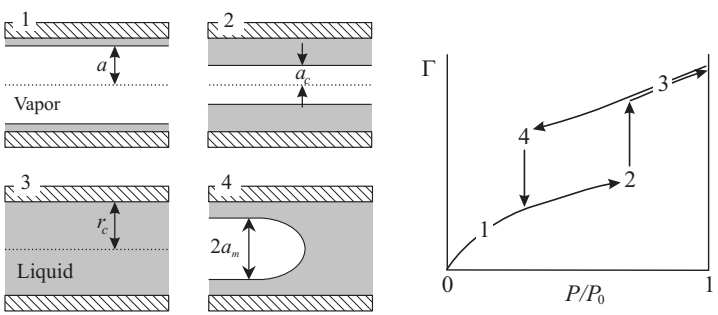



**Figure 9.14:** Filling and emptying of a cylindrical solid pore with liquid from its vapor and the the corresponding schematic adsorption/desorption isotherms.

In the second mechanism the topology of the pore network plays a role [394]. During the desorption process, vaporization can occur only from pores that have access to the vapor phase, and not from pores that are surrounded by other liquid-filled pores. There is a "pore blocking" effect in which a metastable liquid phase is preserved below the condensation pressure until vaporization occurs in a neighboring pore. Therefore, the relative pressure at which vaporization occurs depends on the size of the pore, the connectivity of the network, and the state of neighboring pores. For a single "ink bottle" pore this is illustrated in Fig. 9.15. The adsorption process is dominated by the radius of the large inner cavity while the desorption process is limited by the smaller neck.

In recent years the understanding of the adsorption of vapors has significantly progressed. This is mainly due to the fact that more materials with defined pore sizes can be made [395,

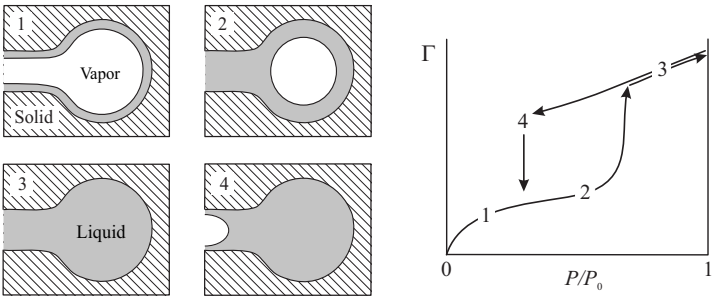



**Figure 9.15:** Filling and emptying of a solid ink bottle pore with liquid from its vapor and the corresponding adsorption/desorption isotherms.

396]. An example of the adsorption to one such material is shown in Fig. 9.16. The siliceous material, called MCM-41, contains cylindrical pores [397]. With increasing pressure first a layer is adsorbed to the surface. Up until a pressure of $P/P_0 \approx 0.45$ is reached, this could be described by a BET adsorption isotherm equation. Then capillary condensation sets in. At a pressure of $P/P_0 \approx 0.75$, all pores are filled. This leads to a very much reduced accessible surface and practically to saturation. When reducing the pressure the pores remain filled until the pressure is reduced to $P/P_0 \approx 0.6$. The hysteresis between adsorption and desorption is obvious. At $P/P_0 \approx 0.45$ all pores are empty and are only coated with roughly a monolayer. Adsorption and desorption isotherms are indistinguishable again below $P/P_0 \approx 0.45$.

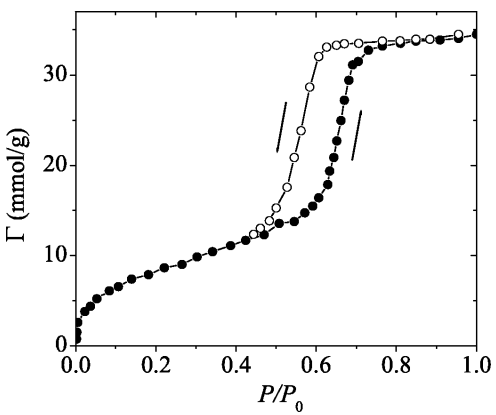



**Figure 9.16:** Adsorption isotherms of Argon at 87 K in siliceous material MCM-41 with cylindrical pores of 6.5 nm diameter. Redrawn after Ref. [398].

## 9.4.4 Special aspects of chemisorption

Chemisorption is usually connected with a chemical reaction. Therefore, it is not surprising that the molecules must overcome an activation energy $E_A$ (Fig. 9.17). Often the molecules

first physisorb to the surface and in a second, much slower, step the bond is established. For desorption both the adsorption energy $Q$ and the activation energy must be overcome. The desorption energy $E_{des}$ is thus larger than the adsorption energy .

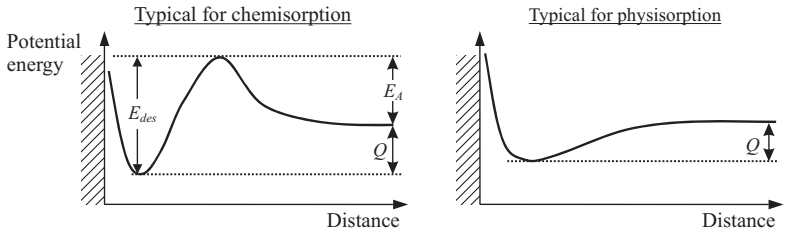



**Figure 9.17:** Potential energy profile versus distance for chemisorption and physisorption.

Experimentally, information about the adsorption and desorption rates is obtained with the help of programmed desorption. One procedure is flash desorption: A surface is instantaneously heated up (normally in vacuum) and we measure the temporal desorption of material, for instance with a mass spectrometer. Heating is usually done with a laser pulse (PLID, pulsed laser induced thermal desorption).

Another possibility is to heat the surface up slowly and to measure the quantity of desorbed material versus the temperature. This is called temperature programmed desorption (TPD) or thermal desorption spectroscopy (TDS). Usually distinct maximums are observed which correspond to the breaking of specific bonds.

For the simple case of a first-order desorption process we briefly describe the analysis; a first order desorption is described by $d\theta/dt = -k_{de}\theta$. Desorption is assumed to be an activated process. The desorption rate, that is the decrease in coverage or the number of molecules coming off the surface per unit time, is

$$k^*_{de} = -\frac{d\theta}{dt} = a\theta \cdot e^{-E_{des}/RT} \tag{9.63}$$

Here, $a \approx 10^{-13}\mathrm{s}^{-1}$ is a frequency factor. In the experiment we increase the temperature of the surface linearly with a rate $\beta$ (in K s$^{-1}$): $T = T_0 + \beta t$. A desorption maximum occurs at a temperature $T_m$. At this maximum we have $dk^*_{de}/dt = 0$. Differentiation of Eq. (9.63) leads to

$$\begin{aligned} \frac{dk^*_{de}}{dt} &= a \cdot \frac{d\theta}{dt} \cdot e^{-E_{des}/RT_m} + a\theta \cdot e^{-E_{des}/RT_m} \cdot \frac{E_{des}\beta}{RT_m^2} = 0 \\ &\Rightarrow \quad a \cdot e^{-E_{des}/RT_m} = \frac{E_{des}\beta}{RT_m^2} \end{aligned} \tag{9.64}$$

Thus, the desorption energy can be calculated from $T_m$ and $a$. An example, the desorption of thiols from gold, is described in Ref. [402].

## 9.5 Adsorption from solution

In this section we consider the adsorption of molecules to surfaces in liquids. Adsorption from solution is a diverse subject. Here, we can only introduce some basic, general features.

A more comprehensive introduction is Ref. [399]. We restrict ourselves to uncharged species and dilute solutions (not binary mixtures). The important subject of polymer adsorption is described in Ref. [400]. Adsorption of surfactants is discussed in Ref. [401]. Adsorption of ions and formation of surface charges was treated in Chapter 5. In dilute solutions there is no problem in positioning the Gibbs dividing plane, and the analytical surface access is equal to the thermodynamic one, as occurs in the Gibbs equation. For a thorough introduction into this important field of interface science see Ref. [8].

Some basic concepts developed for vapor adsorption can also be applied to the adsorption from solution. However, there are also differences. In solution, adsorption is always an exchange process. It is an exchange process in two ways: First, a molecule adsorbing to a surface has to replace solvent molecules. Second, the adsorbing molecule gives up part of its solvent environment. This has several practical consequences:

- Molecules do not only adsorb because they are attracted by the surfaces but because the solution might reject them. An example of this is hydrophobic substances in water. They readily adsorb to many surfaces because of their dislike for water, rather than a strong interaction with the adsorbent.

- A thermodynamic treatment has to take into account the exchange character of adsorption. If one adsorbing molecule replaces $\nu$ solvent molecules at the surface, the whole reaction is

$$A^L + \nu S^\sigma \rightleftharpoons A^\sigma + \nu S^L \tag{9.65}$$

  Here, $A^L$ is the adsorpt in solution, $A^\sigma$ is the adsorbate at the surface, $S^L$ represents a solvent molecule surrounded by other solvent molecules, while $S^\sigma$ is a solvent molecule at the surface.

- Multilayer formation is less common in solutions than in the gas phase because the interaction with the adsorbent is screened by the solvent and the adsorbing molecules have alternative partners.

As an example of adsorption from solution Fig. 9.18 shows the isotherm of $n$-docosane ($C_{22}H_{46}$) and $n$-$C_{28}H_{58}$ adsorbing to graphite in $n$-heptane ($C_7H_{16}$). The longer-chain alkanes are strongly preferred to heptane, indicating that they adsorb in a flat position. A sigmoidal shape is observed with $C_{22}H_{46}$, a hint that the molecules also interact laterally with neighbors. For $C_{28}H_{58}$, adsorption is so strong that even trace amounts all adsorb to the surface. This strong adsorption leads to a high affinity adsorption isotherm (type G in Fig. 9.2). The conclusions — parallel orientation and cooperativity between the adsorbed molecules — are supported by structural studies with the STM.

## 9.6 Summary

- Adsorption isotherms show the amount of a substance adsorbed versus the partial pressure or the concentration.

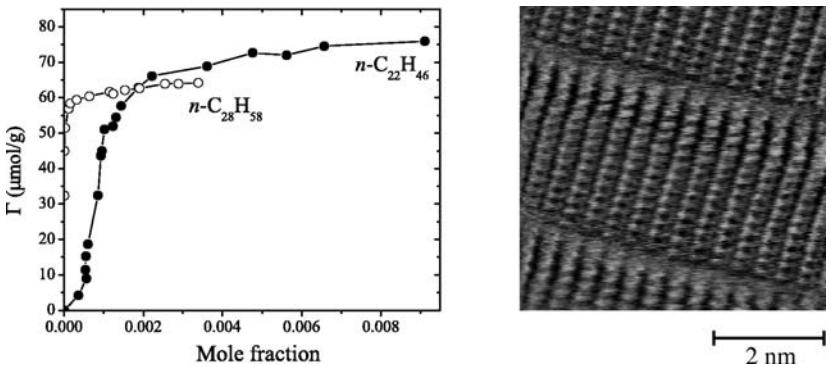



**Figure 9.18:** Left: Adsorption isotherms of long-chain $n$-alkanes from $n$-heptane to graphite (Vulcan 3G). Redrawn from Ref. [403]. Right: Scanning tunneling microscope image of heptacosane ($C_{27}H_{56}$) adsorbed to the basal plane of graphite form an organic solution. The alkanes lay flat on the graphite surface and are highly oriented. Images were kindly provided by J. Rabe [404].

- Physisorption is characterized by a relatively weak binding with adsorption energies of a few 10 kJ/mol. In chemisorption a chemical bond is established between the adsorbate and the adsorbent; desorption energies are in the range of 100–400 kJ/mol.

- To describe adsorption, Langmuir assumed that independent binding sites exist on the adsorbent. This leads to the adsorption equation

$$\theta = \frac{K_L P}{1 + K_L P}$$

  The adsorption isotherm saturates at high partial pressure and the maximal amount adsorbed is that of a monolayer.

- Usually adsorption is more realistically described by the BET model. BET theory accounts for multilayer adsorption. The adsorption isotherm goes to infinity at relative partial pressures close to one, which corresponds to condensation.

- Experimentally, adsorption isotherms are determined by gravimetric or volumetric measurements for powders or porous adsorbents. For isolated flat surface a quartz microbalance or an ellipsometer can be applied to measure adsorption isotherms.

- For porous adsorbents capillary condensation is a common phenomenon. It leads to hysteresis in the adsorption curve.

- In liquids, adsorption is an exchange process in which adsorbed molecules replace liquid molecules.

## 9.7 Exercises

1. Discuss the significance of the Langmuir constant $K_L$. Therefore discuss how the Langmuir isotherm increases for low pressures. At which pressure is the adsorbent half covered?

2. Estimate the Langmuir parameter $K_L$ for nitrogen at 79 K assuming $\sigma_A = 16$ Å$^2$, $\tau_0 = 10^{-12}$ s, $M = 28$ g/mol and $Q = 2$ kcal/mol.

3. BET adsorption isotherm: Verify that $\frac{a_i k_{ad}^1}{a_1 k_{ad}^i} \cdot e^{(Q_1 - Q_\nu)/RT} \approx e^{(Q_1 - Q_\nu)/RT}$ in Eq. (9.38).

4. What is the hydraulic radius for a powder of close-packed spherical particles of radius $R$? 74% of the volume is filled by the particles.

5. Kern and Findenegg measured the adsorption of $n$-docosane ($C_{22}H_{46}$) in heptane solution to graphite [403]. They used a porous graphite with a specific surface area of 68 m$^2$g$^{-1}$ as determined from BET adsorption isotherms with $N_2$. $\Gamma_{max}$, which is assumed to correspond to monolayer coverage, is found to be 88.9 $\mu$mol/g. Can you conclude something about the structure of the adsorbed molecules? What is the area occupied by one molecule compared to its size?

6. Behm et al. studied the adsorption of CO to Pd(100) [323]. Therefore they did a series of TDS experiments at a coverage of $\theta = 0.15$. Using the following heating rates $\beta$ they observed a desorption peak at

| $\beta$ (Ks$^{-1}$) | 0.5 | 1.2 | 2.5 | 4.9 | 8.6 | 15.4 | 25 |
|---|---|---|---|---|---|---|---|
| $T_m$ (K) | 449 | 457 | 465 | 473 | 483 | 489 | 492 |

Calculate the desorption energy $E_{des}$ and the frequency factor $a$ assuming first-order kinetics. Therefore rewrite Eq. (9.64) so that on one side you have $\ln\left(T_m^2/\beta\right)$.

# 10 Surface modification

## 10.1 Introduction

In this chapter we discuss how solid surfaces can be modified. Surface modification is essential for many applications, for example, to reduce friction and wear, to make implants biocompatible, or to coat sensors [405, 406]. Solid surfaces can be changed by various means such as adsorption, thin film deposition, chemical reactions, or removal of material. Some of these topics have already been discussed, for example in the chapter on adsorption. Therefore, we focus on the remaining methods. Even then we can only give examples because there are so many different techniques reflecting diverse applications in different communities.

Before we focus on chemical vapor deposition let us briefly mention other methods for thin film deposition. Films with thicknesses from 1 $\mu$m to several 10 $\mu$m play a fundamental role in everyday life, for example as paints and coatings. Processes during film formation are complex and a discussion would exceed the scope of this book. Introductions are Refs. [407, 408].

Different techniques to deposit thin layers of material have already been described: spin and dip coating in Section 7.5.2, evaporation, sputtering, and molecular beam epitaxy in Section 8.3. Film thicknesses are in the 20 nm to 10 $\mu$m range.

The last three of them require a vacuum for reasons of purity and minimization of secondary reactions between gas molecules. Chemical reactions between gas molecules occur if they have the chance to collide at an energy high enough to overcome the activation barrier. The probability of colliding with each other in the gas phase is a function of the mean free path $\lambda$, which depends on the pressure. The lower the pressure, the higher is the mean free path length and the higher is the technical requirement and the financial expense. Another measure for vacuum quality is the covering time $\tau$ necessary to cover a surface with a monomolecular layer or monolayer of gas molecules. Different types of vacuum are defined in Table 10.1, together with $\tau$ and the mean free path length. For high and ultrahigh vacuum, $\lambda$ becomes larger than the typical size of the chamber.

A combination of evaporation and sputtering is **ion plating**. There, the evaporated metal atoms like titanium are ionized by electrons or plasma ions. A negative applied potential accelerates the metal ions towards the substrate. By adding a reaction gas like nitrogen, a titanium nitride layer is formed on the solid support. This layer is harder and chemically more stable compared to other plating methods. It is an ideal support for adsorbing a new layer with a different material such as gold. This is, for example, used for gold plating in the jewelry industry. A detailed review on ion plating of titanium nitride is Ref. [409].

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

**Table 10.1:** Dependence of the typical mean free path $\lambda$ and time for monolayer coverage $\tau$ (assuming a sticking coefficient of one) on pressure $P$ for nitrogen at 20°C.

| | Pressure $P$ (mbar) | Mean free path $\lambda$ | Covering time $\tau$ |
|---|---|---|---|
| Rough vacuum | 1000 – 1 | 60 nm – 60 $\mu$m | 3 ns – 3 $\mu$s |
| Fine vacuum | 1 – $10^{-3}$ | 60 $\mu$m – 6 cm | 3 $\mu$s – 3 ms |
| High vacuum | $10^{-3}$ – $10^{-7}$ | 6 cm – 600 m | 3 ms – 30 s |
| Ultrahigh vacuum | $< 10^{-7}$ | $> 600$ m | $> 30$ s |

## 10.2 Chemical vapor deposition

In chemical vapor deposition (CVD) complex shaped surfaces can be coated with homogeneous layers, especially when carried out at low pressure (LPCVD, low pressure chemical vapor deposition) (review: Ref. [410]). A gas reacts with the heated substrate surface to give a solid coating and gaseous by-products which have to be removed continously. Layer thicknesses created by chemical vapor deposition are usually in the order 5–10 $\mu$m. In cases where it is necessary to keep the temperature low, a plasma can stimulate the surface reaction in plasma enhanced chemical vapor deposition (PECVD).

■ **Example 10.1.** To deposit $SiO_2$ on solid substrates often TEOS (Tetraethylorthosilicate, $(C_2H_5O)_4Si$) is used as a starting gaseous material. The chemical vapor deposition with TEOS is carried out at about 700°C and 30 Pa, where the mean free path is some 100 $\mu$m. To decrease the process temperature, aggressive gases such as ozone ($O_3$) can be added in minor amounts. Ozone decomposes at about 200°C to oxygen radicals which activate TEOS for deposition. In this way, the process parameters can also be tuned by chemical means [411, 412].

Chemical vapor deposition proceeds in three steps: (1) Transport of educts to the surface region; (2) surface reaction; and (3) transport of the by-products away from the surface. To accelerate the reaction we could initiate a vigorous flow of the gas. The last few microns, however, are usually not influenced by macroscopic flow. Therefore, the molecules have to diffuse to the substrate. Once the approaching molecules have adsorbed to the surface they migrate laterally until they either bind in a potential minimum and form a chemical bond or desorb again. Therefore, the deposition quality is strongly influenced by three processes: migration, a possible nucleation to two-dimensional aggregates, and desorption.

Chemical vapor deposition can even be used to grow epitaxial layers. This method is applied if the solid substrate is a single crystal and the deposition layer adopts a commensurate crystal structure. Epitaxy is in principle a chemical vapor deposition method, carried out at high temperatures to allow for high migration to optimize the crystal structure of the deposition layer according to the underlying template. If the material deposited is the same as the template, the method is called homo-epitaxy, if it is different it is called hetero-epitaxy. Homo-epitaxy is largely used for silicon chips in computer technology. Gases used are silane ($SiH_4$) and chlorosilanes such as $SiCl_4$, $SiHCl_3$, or $SiH_2Cl_2$ applied at temperatures above 1000°C. The compounds decompose to solid silicon coating the surface and chlorine gas

($Cl_2$) or hydrogen chloride (HCl). The silane gas has to be diluted with an inert gas to prevent precipitation in the gas phase. In addition, small amounts of phosphane $PH_3$ or diborane $B_2H_6$ can be added to dope the deposited silicon layer.

Hetero-epitaxy can be used for materials with a similar lattice constant as the template. For electronic devices, silicon on sapphire is often used, because sapphire has a much higher electronic band gap than silicon. This allows us to use silicon chips at high temperature. For the semiconductor gallium arsenide (GaAs) a modified method is applied because of the slightly different lattice constants of silicon and gallium arsenide (0.543 nm for silicon compared to 0.565 nm for gallium arsenide). In this case the method of choice is often metal organic chemical vapor deposition (MOCVD), where trimethylgallium ($Ga(CH_3)_3$) and arsine ($AsH_3$) are used at special deposition conditions.

■ **Example 10.2.** CVD diamond films (Fig. 10.1, review: Ref. [413]). Diamond is the hardest material known. It has a low friction coefficient (in the range of 0.05–0.1) and a high thermal conductivity. It is optically transparent with a refractive index of 2.4, and it is chemically relatively inert. Though pure diamond has a band gap of 5.5 eV and is thus a poor semiconductor it can be doped and electrons and holes have a high mobility. Due to its attractive properties, diamond films are used as coatings for aluminum, copper and their alloys as cutting tools and as optical coatings.

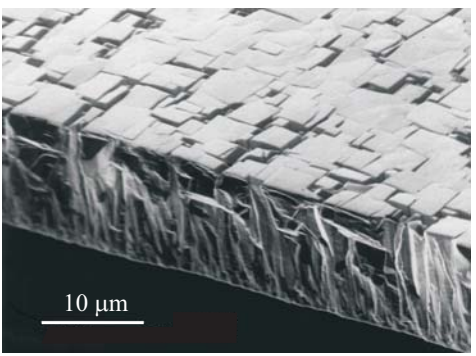



**Figure 10.1:** Scanning electron microscope image of a roughly 10 $\mu$m thick CVD diamond film exposing its (001) facets at the top. The image was kindly provided by X. Jiang [414].

In the standard procedure of CVD diamond film formation the film grows in a chamber containing hydrogen and methane at typically 800°C. At this temperature hydrogen dissociates and forms two radicals: $H_2 \rightarrow 2\,H^*$. The $H^*$ radicals attack methane and methyl radicals are formed: $CH_4 + H^* \rightarrow CH_3^* + H_2$. They also activate carbon on the surface of the growing film by removing surface-bound hydrogen in a similar process: $\sim CH + H^* \rightarrow \sim C^* + H_2$. Now methane radicals can bind to the surface: $\sim C^* + CH_3^* \rightarrow \sim CCH_3$. Cross-links between neighboring carbon atoms are formed after further removal of hydrogen. Please note that there should be no oxygen present in the reaction chamber. How does film growth start on substrates, which do not contain carbon? Either we first put on carbon seeds or a spontaneous nucleation process initiates the formation of small carbon seeds from which the film starts to grow. Unfortunately,

CVD diamond films are not stable on mechanically stressed metals which contain iron, chromium, or nickel and hence steel. In this case films of boron nitride are an alternative.

## 10.3 Soft matter deposition

### 10.3.1 Self-assembled monolayers

This section deals with two types of organic molecules which chemisorb to special solid surfaces: thiols on gold and silanes on silicon oxide (review: Refs. [415–417]). Both consist of one chemical group which binds spontaneously and covalently to the surface and an organic rest group (R). Often a long hydrophobic alkyl chain with $n_C$ methylene units, forming the so-called mesogenic group, is inserted between the chemically reactive and the rest group. Such molecules adsorb as a monomolecular film, a so-called monolayer, on the solid substrate. Since they bind spontaneously such a monolayer is called a "self-assembled monolayer" (SAM). Their final structure on the solid support is determined by three interactions: the strong bond between the head group and the surface, the van der Waals interaction between the mesogenic groups, and the interaction between the tail groups which is often repulsive due to the same orientation of their dipole moments.

Thiols contain a thiol group (SH), also called "mercapto" group, at one end; their general chemical structure is R-SH or $R(CH_2)_{n_C}SH$ if they contain an alkyl chain. Not only thiols but also disulfides ($R_1$-S-S-$R_2$) bind spontaneously to gold and to a lesser degree silver surfaces and form close-packed monolayers (Fig. 10.2) [418, 419]. Forming thiol or disulfide monolayers is practically very simple. If they are volatile they readily bind to any gold surface in a closed vessel. Alternatively, thiols are dissolved in a suitable solvent like ethanol or dichloromethane at a concentration of typically 1 mM. The surface is immersed in this solution for 1–10 h and then it is rinsed to get rid of excess thiols. The binding energy to gold of approximately 120 kJ/mol is relatively strong and the layer is very stable and free of holes. Layers with defined surface properties can be made by selecting an appropriate rest group.

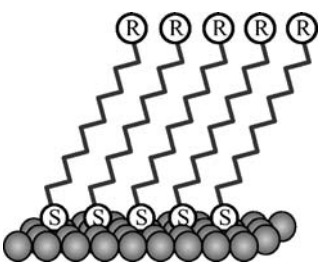



**Figure 10.2:** Schematic of thiols containing alkyl chains on gold(111). Only a row of molecules is shown. In reality a densly packed monolayer is formed.

■ **Example 10.3.** Using thiol monolayers, the wetting properties can be adjusted. Bain et al. [420] coated gold surfaces with monolayers of mixtures of alkanethiols and hydroxy terminated alkanethiols (Fig. 10.3). Alkanethiols result in a high contact angle with water of $\approx 110°$. Pure monolayers of hydroxy terminated alkanethiols show contact angles of $< 20°$. By forming self-assembled monolayers from a mixed solution, surfaces with contact angles between these two values can be formed.

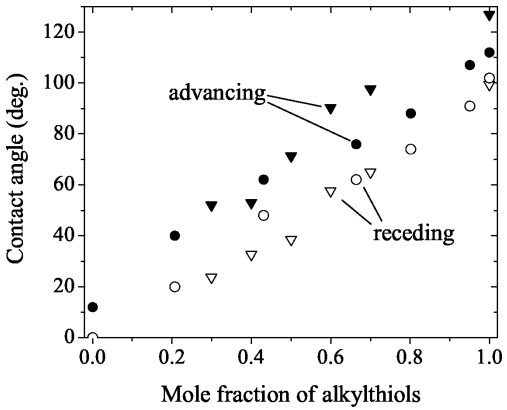



**Figure 10.3:** Change of the advancing (filled symbols) and receding contact angle (open symbols) of a monolayer on gold containing a mixture of $HS(CH_2)_{10}CH_3$ and $HS(CH_2)_{11}OH$. The horizontal axis gives the mole fraction of the alkylthiol with respect to the total thiol concentration in the ethanolic solution from which the monolayers were formed. Redrawn from Refs. [420] (circles) and [421] (triangles).

It is not clear yet, how the chemical bond appears: either thiols and disulfides bind as thiolates, whereby the thiol loses its hydrogen atom and the disulfide dissociates, or both bind as disulfides, whereby two thiols change into disulfides.

■ **Example 10.4.** Long-chain alkanethiols, $CH_3(CH_2)_{n_c}SH$, form a hexagonal $(\sqrt{3} \times \sqrt{3})$ R 30° overlayer on gold(111) (see also Fig. 8.7) with a spacing of 5.0 Å. The diameter of an alkyl chain is only 4.2 Å. In a closed-packed structure such a lattice would not be commensurate with the underlying hexagonal gold(111) lattice with a lattice constant of 2.9 Å. Still, the thiols seem to have a preferential binding site on the gold, probably between three neighboring gold atoms. For this reason, they tilt by $\approx 30°$ to get into contact and reduce their van der Waals energy. An AFM image of such an alkanethiol monolayer on gold(111) is shown on the cover. Other thiols form other structures [422] and large thiols usually do not form two-dimensional crystals at all.

Silanes consist of a silicon atom which can have up to three reactive groups $X_i$, plus one organic rest group R (Fig. 10.4). As with thiols, the rest group is often attached via an alkyl chain. As reactive groups usually hydroxy (∼OH), chlorine (∼Cl), methoxy (∼$OCH_3$, often written as ∼OMe), or ethoxy (∼$OCH_2CH_3$, written as ∼OEt) groups are used.

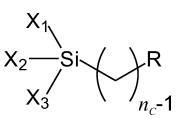



**Figure 10.4:** Structure of a silane.

Silanes react with silanol groups (∼SiOH) on silicon surfaces according to the reaction shown in Fig. 10.5 [423] (review: Ref. [424]). This reaction, which is usually carried out in organic solvents, is called silanization. Trifunctional silanes bind to the surface silanol groups,

releasing 3 HX molecules per silane. This reaction step is called condensation. The double arrow indicates an equilibrium reaction according to the law of mass action. For example, if X is a hydroxyl group, water is set free. It is crucial to exclude water molecules from the solvent solution, as can be already seen from the back arrow in the reaction scheme (Fig. 10.5). Due to the law of mass action less siloxane production occurs, the higher the concentration of water; the equilibrium will be on the left side. Additionally, water will also react with the silanes in solution far away from the surface. The water molecules compete with the hydroxy groups on the silicon surface preferring a fast reaction between the silanes themselves to polysiloxane networks which precipitate. This would spoil the silicon surface because the silanes still react partly with surface OH groups leading to a rough surface.

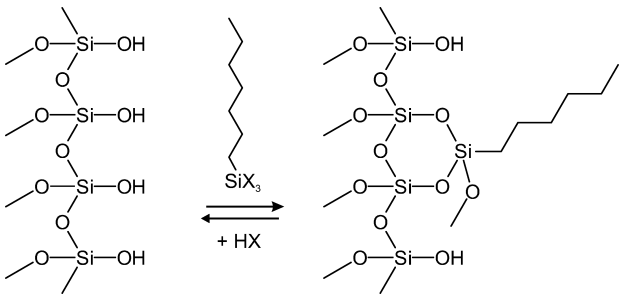



**Figure 10.5:** Silanization reaction scheme for an alkyl silane.

On the other hand, water molecules in traces, as they appear on the hydrophilic silica surface, activate the surface as well as the silanes. This is due to the two-step process of the silanization. When a silane approaches the surface the groups X will react with the water to form silanol SiOH groups, a reaction called hydrolysis. The hydrolysis determines the overall reaction speed and depends on the groups X [425]. It decreases in the series Cl > OMe > OEt. After hydrolysis the silanol condensates in a second step with the silanol groups from the silica surface.

Several competing reactions take place in parallel and determine the quality of the forming SAM. In fact, a stochiometric (1:1) saturation of all silanol groups in the silica surface is sterically impossible. It is more likely that 1 of 5 silanes will bind to the surface whereas the other 4 are cross-linked above it [426]. It resembles more a polysiloxane network with some pinning centers on the surface (Fig. 10.6). Therefore, silane monolayers show a higher roughness and a lower degree of order compared to thiol monolayers on gold.

Practically, the functionalization of silicon is carried out in several steps. First, the substrate is hydrophilized by strong acids like sulfuric acid $H_2SO_4$ mixed with hydrogen peroxide $H_2O_2$ to remove organic residues and/or exposure to a basic mixture of hydrogen peroxide $H_2O_2$ with ammonia to remove metal ions from the silicon surface and create a high hydroxy group density. To remove water except from surface water, the silicon substrates are washed consecutively with different organic solvents which are less and less polar. For example, we can use the sequence methanol, methanol/chloroform 1:1, and finally just chloroform. Afterwards, the silane is dissolved in dry (dry = no dissolved water) organic solvents and the dried substrates are immersed for some hours in the silane solution. Finally, they are taken out and

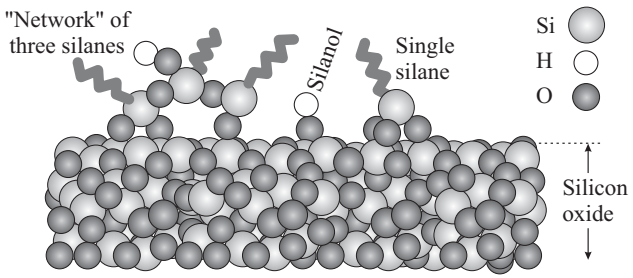



**Figure 10.6:** Schematic drawing of silanes bound to a silicon oxide surface.

washed with the same solvents as before but in the opposite direction becoming more polar to remove physisorbed silanes.

How in general do self-assembled monolayers grow? The quality of the self-assembled monolayers depends on the formation mechanism. Principally, SAM formation includes a diffusion process of the reacting molecule to the surface and an adsorption process at the surface. Depending on the rate of each process, SAM formation is diffusion or adsorption controlled. From the adsorption of a single molecule to the final self-organized film, SAM growth might be thought to happen through different scenarios (Fig. 10.7). Either the molecules adsorb preferentially to already present aggregates on the surface until a closed film is formed (A). Or, first a closed disordered (B) or ordered (C) film is formed and then the molecules within the SAM reorganize and the final film occurs in a second step.

Although thiols on gold and silanes on silicon are the most prominent examples, they are not the only self-assembling systems. Various substrates beside gold and silicon oxide are capable of binding SAMt's such as silver, copper, palladium, platinum and metal oxides as zirconium oxides and indium tin oxide (ITO). As organic molecules alcohols ($\sim$OH), amines ($\sim NH_2$), organic acids ($\sim$COOH), or isocyanides ($\sim$NCO) have been used.

### 10.3.2 Physisorption of Polymers

When small molecules physisorb to a surface it is difficult to get a permanent layer and molecules tend to desorb again. In contrast, macromolecules like polymers usually adsorb irreversibly to surfaces and it is difficult to remove them after adsorption. The reason is that each polymer can bind with many atoms to the surface and even if the binding energy of a single "bond" is below or in the order of $k_B T$, many bonds easily exceed the thermal energy.

A particularly important and well studied case is the adsorption of charged polymers, so-called polyelectrolytes, in an aqueous medium, to solid substrates (review Ref. [427]). They are used industrially as flocculants, for solubilization, as supersorbers, antistatic agents, as oil recovery aids, as gelling agents in nutrition, additives in concrete for building insdustry, or for blood compatibility enhancement amongst other applications [405]. Polyelectrolytes consist of either positively or negatively charged monomers. Common examples are positively charged polyethylenimine, polyallylaminhydrochlorid, polylysin, or negatively charged polystrenesulfonate, and desoxyribonucleic acid (DNA) (Fig. 10.8). The charge arises because of the dissociation of monomer side groups such as amines ($\sim NH_3Cl \rightarrow \sim NH_3^+ + Cl^-$) or

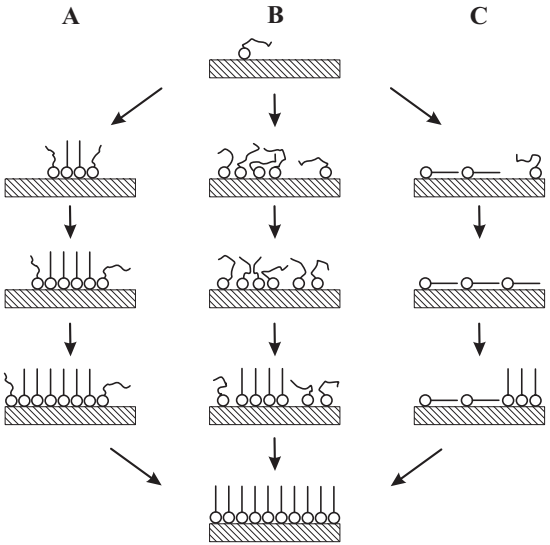



**Figure 10.7:** Schematic of SAM growth. **A**: SAM growth from a two-dimensional gas-phase through a coexistence between gas and a two-dimensional crystalline phase. Alternatively, an intermediate disordered phase might occur after coexistence with a two-dimensional gas (**B**). The final SAM is forming through a coexistence region between the disordered and the finally ordered phase. The intermediate phase might also be ordered (**C**) and reorients to the final SAM through a coexistence between the two ordered phases.

sulfonates ($\sim SO_3M \rightarrow \sim SO_3^- + M^+$). The degree of dissociation determines how many of the monomers in a polyelectrolyte are charged. The more groups are dissociated, the higher the charge of the polymer.

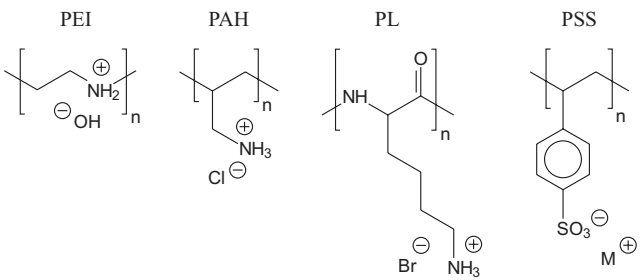



**Figure 10.8:** Chemical structures of positively charged polyethyleneimine (PEI), polyallylaminhydrochlorid (PAH), polylysine hydrobromide (PL) and negatively charged polystyrene-sulfonate (PSS). $M^+$ denotes a metal ion such as $Na^+$.

For polyelectrolytes, the salt concentration has a strong influence on the molecular conformation and the adsorption. The charges on the polymer repel each other electrostatically and in this way tend to elongate the chain. This elongation is reversible: if a salt is added, the charges on the polymer chain are screened and the conformation becomes more compact (Fig. 10.9). Addition of salt has a similar effect as improving the quality of the solvent (see Section 6.7). As a result they tend to adsorb in a more bulky conformation.

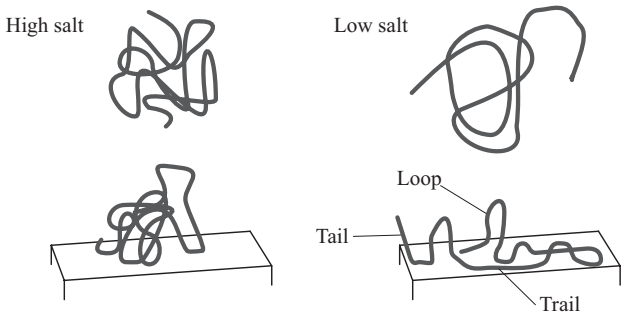



**Figure 10.9:** Conformation of a linear polyelectrolyte depending on salt concentration. At high salt concentration they tend to from a dense random coil which changes to a more stretched conformation at low salt concentration. In addition, the conformation in the adsorbed state is indicated.

In an aqueous medium, polyelectrolytes are subject to a strong adsorption to charged solid substrates each as silicon, due to the electrostatic interaction. The surface potential of silicon wafers, which are oxidized at their surface, is negative and of the order of some $-10$ mV. Thus, a positively charged polyelectrolyte as PEI or PAH adsorbs, rendering the silicon surface positive. Subsequently, a negatively charged polyelectrolyte can be adsorbed on top, then a positively charged again and so on. This layer-by-layer adsorption leads to a polyelectrolyte multilayer (Fig. 10.10) and thus a film on the solid substrate [428]. Experimentally, the adsorption steps are carried out simply by immersing the substrate for some minutes into a beaker containing the aqueous polyelectrolyte solution. The polymer spontaneously adsorbs to the solid surface by self-assembly, i.e. the Gibbs free adsorption energy is negative.

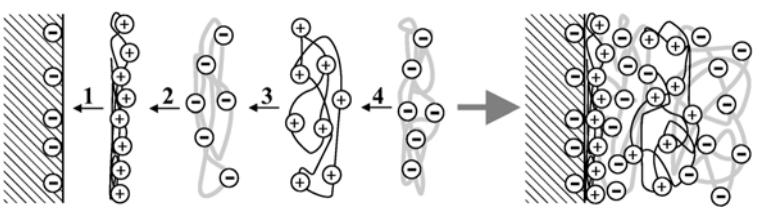



**Figure 10.10:** Layer-by-layer self-assembly of polyelectrolytes on a negatively charged silicon surface.

The thickness of each layer can be influenced by the salt concentration in solution. The higher the salt concentration, the more effective the charges on the polyelectrolyte are screened

and the thicker becomes the adsorbed layer. Usually polyelectrolytes form loops and tails at the surface instead of adsorbing completely flat and well-aligned. For the same reason they strongly interpenetrate adjacent layers. In fact, the effect of the interpenetration can be observed over 2.5 layers. This leads to a high interlayer roughness [429].

Until now we have suggested that the layer structure evolves because of the electrostatic interaction between the oppositely charged polymers. If we assume that each positive charge from one layer finds a negative binding site in the adjacent layer (1:1 complexation), the electrostatic interaction would be screened after one double layer and further adsorption would occur solely due to van der Waals or hydrophobic interactions. Actually, we observe an overcompensation of the charge during the adsorption process. It has been measured that only about one third of the charges in the terminating polyelectrolyte layer are complexed with charges from the underlying layer [430]. The rest of the charges have to be screened by counterions or the charges from the next adsorbing layer.

The variety of materials which can be incorporated into polyelectrolyte multilayers makes them attractive to use as biosensors [431]. Polyelectrolyte multilayers can also be formed on curved surfaces of small particles [432]. After adsorption, the core particle can be chemically dissolved and a hollow polyelectrolyte capsule remains. These capsules are selectively permeable for small molecules like water or certain dyes. The permeability can be tuned externally by varying the ion strength, pH, temperature and solvent nature [433–435]. Therefore, it has been suggested to use them as selective membranes for separation, as well as a possible drug delivery system. The adjustment of their size and permeability allows us to exploit them as micro- or nanocontainers for chemical synthesis and crystallization.

### 10.3.3 Polymerization on surfaces

In the preceeding section we discussed physisorbed polymers. Now we concentrate on chemisorbed polymer layers (review: Ref. [424], see also Section 6.7). Chemisorbed polymers on solid surfaces have the advantage of forming thick flexible layers up to several 100 nm thickness. Due to the flexibility of the polymer chains the layer is relativley homogeneous. Additionally, the large variety of the monomers suitable for surface polymerization leads to a large variety in the surface properties. Also, the mechanical flexibility can be manipulated by the polymer chain density. A high density leads to polymer brushes.

In principle, there are two ways to chemically bind polymers to a solid surface:“grafting to” and “grafting from”. Grafting to means the strong adsorption or chemical coupling of a polymer coil to the surface from solution. This simple approach has the disadvantage of a limited grafting density because already bound polymer molecules hinder other polymers to reach the surface. After coupling to the surface the polymer tries to maintain — as in solution — an entropical given bulky random coil conformation. Thus, it hides binding sites. For higher grafting densities the adsorbed polymer chains have to be stretched at the cost of entropy. This leads to a self-limitation of the density of grafted polymers.

In grafting from this limitation is overcome by polymerizing the polymer directly on the surface. The polymerization is carried out step-by-step after initiation by eligible chemical groups which are called initiators I. The scheme for a grafting-from polymerization is shown in Fig. 10.11. It is also called surface-initiated polymerization or SIP.

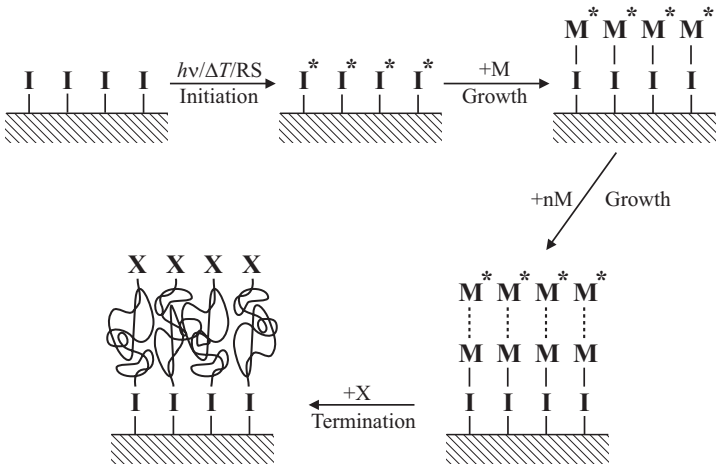



**Figure 10.11:** Scheme for grafting-from polymerization. I is an initiator, M a monomer and X an end-group for further coupling reactions.

Generally, for polymerization, three kinetic steps have to be considered, regardless of their occurence in solution or at a surface (for a standard textbook on macromolecules, see Ref. [436]). At first the polymerization has to be started or initiated. For the grafting-from technique initiator molecules have to be immobilized at the solid surface. Usually, the initiator is a very reactive species which is activated by light or temperature changes or by another reactive chemical added to solution. Therefore, chemical tricks have to be applied to attach the initiator at the surface without already activating it by accident. For surface immobilization SAMs can be used as described in Section 10.3.1 [437]. After surface immobilization the activation leads to a radical, a positively or a negatively charged initiator group, denoted by a star in Fig. 10.11. By reaction with a monomer M the activated site is transferred to the monomer which itself reacts with another monomer, etc. This chain growth is terminated (desactivation) if all of the monomers are used or side-reactions, especially for radical polymerization, stop the reaction. Such side-reactions can be e.g. a combination of adjacent growing radicals to one desactivated chain or a transfer of the reactive site to a neighbor chain, a solvent molecule, or a monomer [438].

Depending on the type of activation, the reaction is called free radical, living anionic, or living cationic surface reaction. In free radical polymerization the reactive site is an electron, in living anionic polymerization, it is a negative charge and in living cationic polymerization it is a positive charge. The term “living” denotes a pronounced difference between radical and the ionic reactions: the monomer can be added step-by-step by the experimenter. This leads to a more unique chain length compared to a radical polymerization. Additionally, radical reactions promote side-reactions as crosslinking or a combination between polymer chains, leading to a higher polydispersity. Moreover, the initiator decomposes throughout the propagation of the reaction and the monomer transfer to these reactive grafting sites becomes more difficult with increasing polymer brush density. In fact, at the end of the radical surface

polymerization a Trommsdorf-like effect has been observed, known from bulk. There, the viscosity in the surface is so high that the reaction heat cannot be exchanged with the bulk phase anymore. In bulk this effect might lead to explosions. These disadvantages are overcome by living ionic polymerizations. Unfortunately, they are very sensitive to contamination which is a crucial point at surfaces. Therefore, ionic polymerization is more difficult and technically demanding.

The need to better control surface-initiated polymerization recently led to the development of controlled radical polymerization techniques. The trick is to keep the concentration of free radicals low in order to decrease the number of side reactions. This is achieved by introducing a dormant species in equilibrium with the active free radical. Important reactions are the living radical polymerization with 2,2,4,4-methylpiperidine N-oxide (TEMPO) [439], reversible addition fragment chain transfer (RAFT) which utilizes so-called iniferters (a word formed from *ini*tiator, chain trans*fer* and *ter*minator) [440], and atom transfer radical polymerization (ATRP) [441–443]. The latter forms radicals by added metal complexes as copper halogenides which exhibit reversible reduction–oxidation processes.

## 10.4 Etching techniques

A conceptually simple way of changing surface properties is to remove parts of it. These methods are summarized as "etching techniques" [407, 444]. In particular in microfabrication etching techniques are essential. Etching techniques are characterized by several properties:

- Wet and dry etching. Wet etching is done in solution containing aggressive chemicals, dry etching is done by gas molecules
- Etching can be physically or chemically. Physical etching denotes an etching process with inert ions such as argon ($Ar^+$), whereas chemical etching uses reactive ions such as the oxygen ion ($O^+$) or the chlorinetrifluoride ion ($ClF_3^+$), neutral gases as xenondifluoride ($XeF_2$) or radicals, i.e. atoms or molecules with single unpaired electrons, as highly reactive atomic fluorine ($F^*$), oxygen ($O^*$), or trifluormethyl radicals ($CF_3^*$).
- Surface etching can be isotropic or anisotropic. For isotropic etching the etching rate is the same in any direction, for anisotropic it is higher in one direction.
- Selectivity refers to the material which is actually removed in a multicomponent system.

Often, we want to achieve a high selectivity and a high anisotropy. Practically these exclude each other and we have to compromize. To illustrate this we show different possible combinations for dry and wet etching below, leading to different anisotropy and selectivity. Figure 10.12 shows a bilayer system with different materials to which a certain selectivity must be applied. Such layered systems are etched through a mask to inscribe a certain desired three-dimensional structure for which a certain anisotropy must be maintained.

Several procedures have been developed for dry etching. They are listed in the following from low to high selectivity accompanied by increasing isotropy.

- **Ion beam etching** (IBE) is a physical etching method in which inert ions are accelerated with an ion gun, towards the substrate. The substrate itself is placed in a separate etching

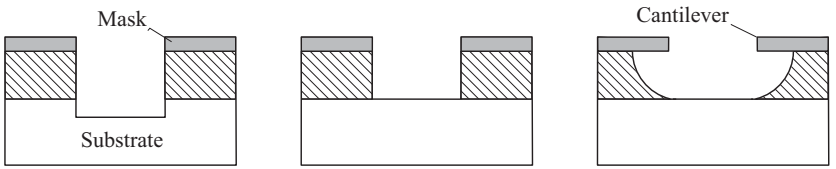



**Figure 10.12:** A microstructure after anisotropic, non-selective etching (left), anisotropic, selective etching (middle), and isotropic, selective etching (right).

chamber in high vacuum. Ion beam etching is relatively anisotropic, but has a poor selectivity.

- **Sputter or ion etching** (IE) is also a physical etching method with inert ions (such as $Ar^+$) from a plasma which are accelerated towards the substrate. In this case the substrate is in contact with the plasma. The etching profile is anisotropic, the selectivity is poor.
- **Reactive ion beam etching** (RIBE) is similar to ion beam etching but with reactive ions. The etching profile depends on the special conditions, the selectivity is lower than in ion beam etching.
- **Reactive ion etching** (RIE) is like ion etching but with reactive ions.
- **Plasma etching**. (PE) Physical–chemical etching with free radicals, supported by ions. Etching profile anisotropic–isotropic, good selectivity.
- **Barrel etching**. (BE) Chemical etching with free radicals only. Etching profile highly isotropic, selectivity good. Due to the isotropic etching, cantilevers can be formed (Fig. 10.12, right.)

The technical equipment has to be engineered according to the special needs of selectivity and anisotropy. In ion etching, for example, two electrodes of different size are fixed as a plate capacitor in an evacuated chamber. The chamber is filled with argon at a pressure of 0.5 to 10 Pa. The substrate to be etched is placed on the smaller electrode. The larger electrode is grounded. A voltage of 0.1–1 kV is applied. Under these conditions a gas discharge occurs between the electrodes. Electrons are accelerated towards the positively charged electrode. They release their energy during collisions with neutral argon atoms, causing light emission, which can be seen as a discharge glow, or ionization. Consecutively, the ionized argon atoms are accelerated towards the negatively charged electrode. There, they might generate more secondary electrons when colliding with the electrode surface or the ions are neutralized again with electrons (recombination). In equilibrium, the rates for ionization and recombination are the same.

To prevent charging of insulating electrodes, an ac voltage is applied. The charge accumulation on the electrodes during one half-cycle is neutralized by the opposite charges during the next half-cycle. Usually, a radio frequency (rf) is used to generate the gas discharge. A plasma in contact with a surface is forming a dark space close to it. Since the lighter electrons are much more mobile than the heavy argon ions, those accelerate faster towards the electrodes.

But a charging of the electrodes is inhibited by a negative charge built-up which is attracting argon ions. The electrical potential at the smaller electrode, where the substrate is located, compared to the potential at the grounded and bigger electrode, is inversely proportional to the area ratios of the electrodes [444]. Therefore, this electrode exhibits a higher negative potential than the bigger electrode and the bombardement with the $Ar^+$ ions is much more intensive and perpendicular to the surface. Thus, it acts less selective but highly anisotropic.

We have already learned that the free path length of gas ions is large compared to the structures to be etched if the gas pressure is sufficiently low. Also, the thermal energy is low compared to the energy gained in the electric field. As a consequence, the ions move practically perpendicular to the surface and the etching process is highly anisotropic in ion etching. In fact, the etching rate depends on the angle under which the ions collide with the surface. For etching in normal direction a momentum is transferred vertically to the surface lattice. To remove an atom from its lattice site the momentum transfer should be maximal; the direction of the momentum vector changes by 180°. Decreasing the incident angle means less change of momentum. On the other hand an ion travelling parallel to the surface has a longer reaction time because it spends more time at the surface. As a result, there is an angle, different from 90° to the surface at which the etching rate is a maximum. This often leads to oblique edge angles at longer etching times.

Pure chemical barrel etching is in principle very similar. Usually, it is carried out in a tubular reactor containing the gas at about 10 to 100 Pa. An alternating voltage of typically 1 kV is applied. The trick: around the sample in the middle of the tube there is a shielding tube with holes in it. This allows the formed radicals to diffuse to the sample surface through the holes. The sample itself has no contact with the discharged gas. The diffusion is isotropic, so is the etching, but due to the special chemical species involved barrel etching acts selectively to certain surfaces only. Therefore, it is useful and used to remove special layers as photoresists, i.e. light sensitive coatings, completely.

Reactive ion etching and plasma etching are carried out in the same device like ion etching but the inert gas argon is replaced by reactive gases such that the etching process becomes more isotropically. For plasma etching [446] the substrate to be etched is placed on the larger electrode under higher pressure (10 – 100 Pas), leading to lower energy transfer onto the surface. Therefore, plasma etching is acting by far more isotropic but chemically selectively. For reactive ion etching the substrate is placed on the smaller electrode under lower pressure as in ion etching. Now, the action on the surface is harsher again and reactive ion etching is less selective but, due to the smaller electrode, more anisotropic compared to plasma etching. Since in both methods there is no discrimination of inert or reactive ions, they are not as selective and isotropic as barrel etching.

■ **Example 10.5.** When applied to soft hydrophobic polymeric surfaces such as the surface of polystyrene, methods like plasma and reactive ion etching lead to a hydrophilization of the surface. This is due to a chemical reaction with oxygen which imposes polar groups into the polymer surface. These polar groups "like" water and the contact angle of water drops is small. Unfortunately, hydrophilisation is not permanent and the contact angle increases with time again. Explanation: Surfaces with hydrophilic groups have a higher energy than hydrophobic ones. In polymers these groups are still mobile and they tend to move into the polymer which becomes more hydrophobic and thus low-energetic again [447, 448].

Dry etching techniques have one common disadvantage: they are more useful for small etching depths in 1 $\mu$m-range and small areas. If we have to etch deep structures out of large areas like silicon wafers, wet chemical procedures have to be applied. Isotropic etching for silicon surfaces is based on an oxidation to a thick oxid layer and successive dissolution of this oxid layer. The typical reactive mixture is made of hydrofluoric acid (HF) and nitric acid ($HNO_3$). It is assumed that $HNO_3$ oxidizes the silicon to $SiO_2$ which is destroyed by HF according to

$$\mathrm{SiO_2 + 6\,HF \rightarrow SiF_6^{2-} + 2\,H_2O + 2\,H^+} \tag{10.1}$$

Therefore, fluoric acid should never be filled into glass containers! Be careful while handling fluoric acid. It is a strong acid, very toxic, and causes severe burns.

Anisotropic wet etching is the key technique for the fabrication of micromechanical structures as microchannels, microreactors but also mobile parts like micro-toothed wheels etc. The etching rate in reactive chemicals' solution strongly depends on the lattice orientation of the single crystal exposed to the solution and might differ by several orders of magnitude. The orientation can be varied by the angle under which the single crystal is cut into pieces. By this it is possible to etch sharp-edged structures with a V- or U-shaped cross-section into crystalline material.

For silicon we can use basic etching solutions like sodium, potassium, lithium, or ammonium hydroxide (NaOH, KOH, LiOH, $NH_4OH$). Another important group of etching solutions contains organic compounds: a mixture of ethylenediamine ($NH_2CH_2CH_2NH_2$), a base due to the amino groups, pyrocatechol, a complexing agent for silicon, pyrazin or, or a solution of hydrazine $N_2H_4$ in water. An advantage of organic solutions is a smoother finish due to the relatively high etch rate (some 10 $\mu$m/hour at higher temperature), a missing contamination with metal ions, and a better fabrication of undercutting cantilevers (Fig. 10.12, right).

The etching rate also strongly depends on the presence of additional atoms like boron or phosphorus in the silicon lattice, all typical dope materials. They are so effective that they practically stop the etching process. This is called etch-stop. How does the etch-stop process work? Silicon is affected by the basic hydroxyl ions ($OH^-$). They crack the bonds to next neighbor atoms in the silicon surface and, because of their negative charge, the silicon surface becomes negatively charged. This charge is moving in the highest unoccupied energy level of the silicon, the so-called conduction band. Water molecules approaching this activated silicon surface are reduced according to

$$\mathrm{2\,H_2O + 2\,e^- \rightarrow 2\,OH^- + H_2} \tag{10.2}$$

The water molecules pick up electrons and thus are reduced to hydroxyl anions. Eventually two hydroxyl ions and hydrogen gas is generated. The hydroxyl ions crack more silicon bonds until silicon is completely removed by the surrounding liquid phase. If there are boron atoms present they offer unoccupied energy levels just some 10 meV above the highest fully occupied energy level in the silicon lattice, called the valence band. Due to the additional empty energy levels the electrons from the chemical reaction are caught away and cannot further reduce water molecules to hydroxyl ions. The etching comes to an end.

## 10.5 Summary

- To coat surfaces with films of inorganic materials, vacuum (or low pressure) deposition methods are used such as evaporation, sputtering, molecular beam epitaxy, and ion plating. Such films are typically up to a few 10 nm thick. Films of typically 10 $\mu$m thickness are produced by chemical vapor deposition in which a gas reacts with the surface, usually at high temperatures, to give a homogeneous, solid coating.

- Surface properties can drastically be changed by forming self-assembled monolayers without the need for special technical equipment. In this case organic molecules with a special functional group bind spontaneously to the surface and form a dense monolayer. In addition to the functional group they have a rest group attached which determines the surface properties. Prominent examples are silanes binding to silicon oxide and thiols (and disulfides), which form self-assembled monolayers on gold.

- Polymers on surface can change the surface properties drastically. Even physisorbed polymers are usually bound irreversibly due to the large number of binding sites. To produce dense and thick layers, polymers have to be grafted from the surface.

- Surface properties can also be changed by removing material from the surface. This is done in various dry and wet etching processes.

## 10.6 Exercises

1. Hydrophobization/hydrophilization of solid surfaces. a) How would you hydrophobize a silicon surface? b) How would you hydrophilize a gold surface? c) Suggest and explain briefly methods to check the quality of your coatings concerning wettability, thickness, and structure.

2. Protein adsorption on solid substrates. Proteins are macromolecules built up from a pool of 20 amino acids (COOH-CHR-$NH_3$) linked together by peptide bonds ($\sim$CHR$_1$-NH-CO-CHR$_2\sim$) in different combinations and fractions. The aminoacids can be charged or uncharged. The proteins have different functions in the body, e.g., as biocatalysts which accelerate fundamental biochemical reactions. Other functions are, e.g., as structural proteins like in collagen in hairs, as molecular motors like myosines in muscles, as sensors like rhodopsine in eyes etc. Characteristic is a specific three-dimensional structure as the fibrous collagen, a more globular structure as in haemoglobin containing red blood cells or Y-shape structures as in antibodies. It is crucial for the protein to maintain this structure. a) What do you expect happens to the structure of a protein when it adsorbes on a solid surface? Hint: First think about the forces which hold a protein together. b) How can we prevent a protein to adsorb on a surface?

3. One important application of proteins on surfaces is the glucose sensor. It measures the glucose content, e.g., of blood which is crucial for diabetics. Glucose is oxidized to gluconic acid by oxygen with the help of glucose oxidase. Simultaneously, oxygen is reduced to hydrogen peroxide. a) Complete the reaction schemes below. Write down the

# 11 Friction, lubrication, and wear

Friction is the force between interacting surfaces that resists or hinders their relative movement. Wear is defined as the progressive loss of material from a body caused by contact and relative movement of a contacting solid, liquid or gas. The aim of lubrication is to reduce friction between surfaces and minimize wear. Today the research field of friction, lubrication and wear is called "tribology". This word is derived from the ancient Greek word "tribein" (meaning rubbing) and was first used in 1966 in a publication titled, *The Jost Report: Lubrication (Tribology) Education and Research* published by Her Majesty's Stationery Office (HMSO) in the United Kingdom. Since then it became the common term to describe the science of friction, wear, and lubrication. Introductions to this field are Refs. [452, 453].

In spite of the strong economic importance of friction and wear and the resulting scientific effort, our understanding of the fundamental processes is still rudimentary. This results from the complexity of these topics. In addition, this complexity demands a multi-disciplinary approach to tribology. In recent years the development of new experimental methods such as the surface forces apparatus, the atomic force microscope, and the quartz microbalance made it possible to study friction and lubrication at the molecular scale. However, this new wealth of information does not alter the fact, that there are no fundamental equations to describe wear or calculate friction coefficients. Engineers still have to rely largely on their empirical knowledge and their extensive experience.

## 11.1 Friction

### 11.1.1 Introduction

Friction is one of the oldest subjects in the history of science and technology and its importance in everyday life is obvious. Nevertheless, there is no precise macroscopic theory of friction that would allow us to predict the frictional force between two given bodies. For nearly every rule there are exceptions. An introduction to friction phenomena is Ref. [454], a detailed book on sliding friction is Ref. [455].

It is useful to differentiate between static and dynamic friction. Dynamic friction, also called kinetic friction, is the mechanical force between sliding or rolling surfaces that resists the movement. Static friction must be overcome to start the movement between two bodies which are initially at rest.

In this section we discuss dry friction, also called solid or Coulomb friction. Solid friction occurs when two solid surfaces are in direct contact without any other components like lubricants or adsorbed surface layers involved. The reader might object that in practice no such

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

surface exists and that all surfaces are "contaminated". In many cases, the effect of the surface layer can, however, be neglected.

### 11.1.2 Amontons' and Coulomb's Law

The first recorded systematic studies on static friction have been carried out by Leonardo da Vinci.[1] He had already stated that friction does not depend on the contact area and that doubling the weight doubles the friction. The most important empirical law found for describing friction was published in 1699 by Guillaume Amontons.[2] Like da Vinci he measured the force $F_F$ required to slide a body over a solid surface at a given load $F_L$ (Fig. 11.1). The load is usually the weight of the body but it can also contain an additional external force pushing the body down. Amonton found that the frictional force is proportional to the load and does not depend on the contact area. For example, in Fig. 11.1 the loads $F_L^1 = F_L^2$ are equal, then the frictional forces are also equal $F_F^1 = F_F^2$. In other words: the coefficient of friction $\mu$ defined by

$$F_F = \mu F_L \tag{11.1}$$

should be constant and independent of the contact area. Amontons himself gave a value of 1/3 for the coefficient of friction. Although Amontons' law is today used to describe dry friction, Amontons himself employed greased surfaces in his experiments — using conditions which today would be denoted as boundary lubrication (see below).

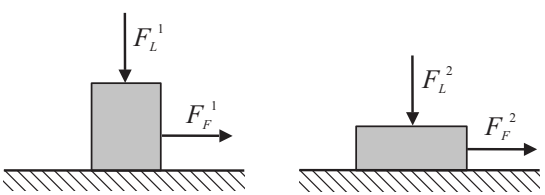



**Figure 11.1:** Amontons' Law of Friction: the frictional Force does not depend on the contact area and is proportional to the load.

Amontons' law is purely empirical and results from the interplay of very complex processes that we are only recently beginning to understand. At first sight, Amontons' finding that the friction force is independent of the surface area is quite surprising. Around 1940, Bowden and Tabor pointed out, that the true area of contact $A_{real}$ between two solids is only a small fraction of the apparent contact area [456]. This is because of surface roughness. Practically almost all surfaces are rough. Some surfaces, such as glass or polished metal, might appear optically smooth. Appearing optically smooth only means that the roughness is significantly below the wavelength of light. On the nanometer scale these surface are still rough. Due to the roughness, the surfaces touch each other only at some microscopic contacts

[1] Leonardo da Vinci, 1452–1519. Italian scientist, inventor, and artist.

[2] Guillaume Amontons, 1663–1705. French army engineer.

(asperities), also referred to as microcontacts or junctions. The friction force is equal to the force necessary to shear these junctions and is therefore given by:

$$F_F = \tau_c \cdot A_{real} \tag{11.2}$$

Here, $\tau_c$ is the yield stress during shear. The true contact area $A_{real}$ depends on the compliance of the materials. For soft, rubberlike materials, the real contact area will be larger than for hard materials like steel (Fig. 11.2).

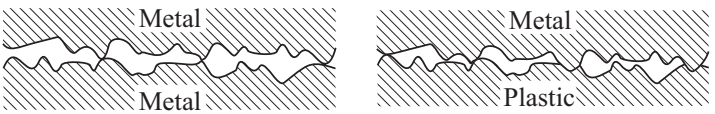



**Figure 11.2:** Surface roughness and true contact between two hard solids and a hard and a soft solid.

How can the actual contact surface be measured? One possibility is to measure the electrical resistance between two conductors and calculate the contact area from the measured resistance and the specific resistivity of the materials. Another possibility is to use an IR sensitive microscope to measure hot spots of a transparent solid that is in contact with a hot surface. With these methods it was found that the friction force is, in fact, proportional to the actual contact area. This implies that the true contact area must increase linearly with load. To illustrate how this is possible, we consider two extreme cases. In the first case, purely elastic deformation is considered. In the second case, we assume plastic deformation of the microcontacts.

- **Elastic deformation**. For small loads we can use the Hertz model as a simple approximation. The microcontacts are thereby assumed to be spherical. Hertz theory predicts an actual contact surface for an individual sphere on a plane (see Eq. 6.64):

  $$A_{real} \propto F_L^{2/3} \tag{11.3}$$

  This does not yet lead to linear dependence between load and friction. Greenwood [457] therefore assumed a gaussian distribution of sphere sizes, as is observed for most engineering surfaces. He showed that under these conditions $A_{real}$ is, in fact, proportional to $F_L$ resulting in a linear dependence between true contact area and normal load.

- **Plastic deformation**. If we bring two surfaces into contact, then the pressure at each micro contact can be very high. If the pressure exceeds the yield stress $P_m$, the microcontacts will deform plastically. The yield stress is the maximum pressure before the material starts to deform plastically. Plastic deformation continues until the increase in contact area causes the contact pressure to drop below the yield stress. Therefore, the actual contact surface $A_{real}$ is considerably affected by the yield stress. Plastic deformation stops when

  $$A_{real} \cdot P_m = F_L \tag{11.4}$$

Therefore, the actual contact area is independent of the apparent contact area. It increases linearly with increasing load. Typical values for the yield stress of metals are $10^8 - 10^9$ Pa. Thus, for a 10 kg block of a metal the true area of contact is of the order of (0.3–1 mm)$^2$.

For both cases, the assumption that friction is proportional to the true contact area $A_{real}$ directly leads to Amontons' law of friction.

The significance of surface roughness for explaining Amontons' law immediately provokes the question of how the friction depends on the surface roughness. Usually, friction of dry and clean surfaces does not change much with roughness, unless the roughness becomes so large that the two sliding bodies move up and down due to macroscopic asperities. There is a second exception. For very smooth and clean surfaces, the friction force may actually become very high. An example is smooth metal surfaces in a vacuum that may exhibit very high coefficients of friction. The reason is the intimate contact between the surfaces that leads to strong adhesion and even cold welding can occur. In lubricated systems, rough surfaces tend to produce higher friction.

Another important empirical law of dry friction is due to Coulomb. Coulomb, who became famous mostly for his studies on electricity and magnetism, made extensive studies on dry friction and formulated a further law: The frictional force between moving surfaces is independent of the relative speed. At first sight Coulomb's law is also surprising. For the movement of a particle in a viscous medium we know that the friction or drag force acting on this particle is in fact *proportional* to the velocity of the particle (Example: Stokes friction for a sphere moving through a fluid). The explanation was given by Tomlinson in 1929 [458] and will be discussed in Section 11.1.9. Usually there is a slight decrease of dry friction at higher speeds. This is related to an increased surface temperature that leads to a reduced shear strength of the microcontacts. In lubricated systems, the dependence of friction on speed becomes more complicated and varies according to the different interaction regimes (see Fig. 11.12 ).

### 11.1.3 Static, kinetic, and stick-slip friction

Static friction is higher than dynamic friction. Therefore, we have to distinguish between a static friction coefficient $\mu_s$, which refers to the force that must be exceeded in order for a motion to start, and a kinetic friction coefficient $\mu_k$, which is related to the force needed to sustain sliding. In general, $\mu_k \leq \mu_s$.

If the dragging force is coupled elastically to a body this may lead to a so-called stick-slip motion [459]. Examples for stick-slip motion are the excitation of a violin string by the bow or the squeaking of brakes or of doors. To illustrate stick-slip motion, we take as a model a block of mass $m$ that is connected via a spring (spring constant $K$) to a hook that is moved with constant velocity $v$ (see Fig. 11.3). If the block is moving with the same speed as the hook and the elongation of the spring is equal to $\Delta x = \mu_k mg/K$, the system is in equilibrium. When we start pulling the hook, with the mass at rest, it will not move until the elongation force of the spring has reached the value $F_{spring} = \mu_s mg$ (= stick phase). Then the block starts to move and is accelerated to a velocity greater than $v$ since the kinetic friction coefficient is smaller than the static one (=slip phase). This rapidly restores the spring to a more relaxed

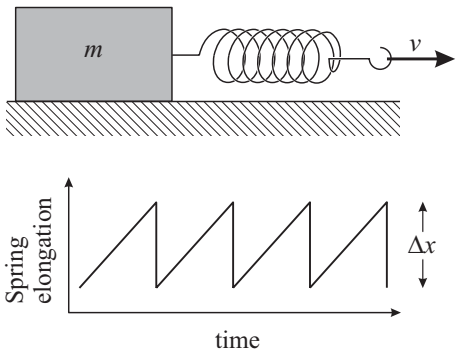



**Figure 11.3:** Example of a system that exhibits stick-slip friction. Stick-slip is also illustrated as a schematic plot of spring elongation versus time at constant pulling speed $v$.

length. The drag force on the block decreases, causing it to come to rest again, and the whole process starts over again.

In technical applications, stick-slip friction is detrimental in terms of wear, vibrations, and precision of movement. Stick-slip:

- is more pronounced at small velocities,
- increases with increasing difference between $\mu_s$ and $\mu_k$,
- is more significant when using soft springs.

To avoid stick-slip, one should try to make the spring constant high enough (using stiff materials and stable constructions). It can be shown, that stick-slip may also arise from the velocity dependence of the friction coefficient [460]. When the friction coefficient decreases with sliding velocity, stick-slip is amplified. When the friction coefficient increases with velocity, stick-slip is damped out. The former is usually the case at low speeds, certainly for the transition from static to dynamic friction, whereas the latter prevails usually at high velocity.

Important examples of stick-slip are earthquakes that have long been recognized as resulting from a stick-slip frictional instability. The use of a full constitutive law of rock friction that takes into account the time dependence of $\mu_s$ and the dependence of $\mu_k$ on speed and sliding distance can account for the rich variety of earthquake phenomena as seismogenesis and seismic coupling, pre- and post-seismic phenomena, and the insensitivity of earthquakes to stress transients [461].

Macroscopic stick-slip motion described above applies to the center of mass movement of the bodies. However, even in situations where the movement of the overall mass is smooth and steady, there may occur local, microscopic stick-slip. This involves the movement of single atoms, molecular groups, or asperities. In fact, such stick-slip events form the basis of microscopic models of friction and are the explanation why the friction force is largely independent of speed (see Section 11.1.9).

### 11.1.4 Rolling friction

Experience tells us that much less force is required to roll a wheel or cylindrical object rather than to slide it. The coefficient of rolling friction is usually defined by

$$M = \mu_r F_L \quad \text{or} \quad F_F = \mu_r \frac{F_L}{R} \tag{11.5}$$

Here, $M$ is the torque of the rolling object, $\mu_R$ is the coefficient of rolling friction, $F_L$ is the normal force (load), $F_F$ is the friction force and $R$ is the radius of the rolling object (see Fig. 11.4). In this case the friction coefficient is not dimensionless but has the unit of length. Other definitions use

$$F_R = \mu_r F_N \tag{11.6}$$

in analogy to Amontons' law, to have a dimensionless coefficient of friction. Typical values for the coefficient of rolling friction are of the order of $10^{-3}$.

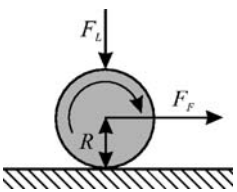



**Figure 11.4:** Sphere or cylinder rolling over a planar surface.

Ideally, that is for infinitely hard solids, a rolling sphere or cylinder makes contact with the underlying surface at only one point or a single line, respectively. In this cases, rolling friction would in fact be zero as there is no relative movement of the contacting surfaces. In real systems, there is always a finite contact area, as we have seen in Section 6.8. As a result, there are different sources of energy dissipation and thus rolling friction:

- Relative sliding (microslip) of the contacting surfaces. When the elastic moduli of the two bodies rolling on each other are different, there will be a different amount of stretching of the two materials in the contact area, leading to a slip. This was already recognized by Reynolds in 1876 [462]. Heathcote [463] showed that, if a sphere is rolling on a soft surface, the instant center of rolling is not identical with the center of the sphere, but lies just above the lowest point of contact. Therefore, the single contact surface elements have different distances from the instant center of rolling and this enforces slip. The contribution of slip to the total rolling friction seems to be low since the dependence of rolling friction on lubrication is very weak [459]. It may nevertheless be important from the engineering point of view to separate fixed contact and slip areas when constructing gear transmissions.

- Adhesion. During rolling there is a continuous generation of new contact area at the front and a continuous disrupting of contact at the back side. Viewed from the point of surface energies, the two effects should just compensate. Usually, however, this is not the case. Depending on the detailed interaction, disruption of material bridges at the back side leads to energy dissipation. An extreme example would be rolling friction on an adhesive tape.

- Plastic deformation. If normal or tangential stresses become too high, plastic deformation of the contact area will occur. Under certain circumstances, this can lead to plastic shearing, parallel to the surface, without plastic penetration normal to the surface.

- Viscoelastic hysteresis. Relaxation processes within the materials can contribute to rolling friction especially in the case of viscoelastic materials.

Rolling friction is often found to be proportional to the velocity, but more complex relationships may be observed, depending on the combination of the bodies. For a soft, viscoelastic sphere on a hard substrate, Brilliantov et al. [464] predicted a linear dependence of rolling friction on speed. For a hard cylinder on a viscous surface, a much more complex behavior was found [465, 466]. At lower speeds, the rolling friction increases with speed to reach a maximum value and then decreases at higher speeds. The reason is an effective stiffening of the substrate at higher speeds.

### 11.1.5 Friction and adhesion

Friction becomes stronger with increasing adhesion between the two solids. Strong adhesion between two bodies is caused by strong attractive forces, for example van der Waals forces. Usually we can take this adhesion force $F_{adh}$ into account by simply adding it to the load. Equation (11.1) is then replaced by

$$F_F = \mu(F_L + F_{adh}) \tag{11.7}$$

For macroscopic objects the adhesion force is often small compared to the load. For microscopic bodies this can be different. The reason is simple: the weight of an object sliding over a surface usually decreases with the third power of its diameter (or another length characterizing its size). The decrease of the actual contact area and hence the adhesion force follows a weaker dependence. For this reason, friction between microbodies is often dominated by adhesion while in the macroscopic world we can often neglect adhesion.

■ **Example 11.1.** The sliding frictional force between a spherical silica particle ($SiO_2$) and a planar silicon wafer increases with the true contact area as calculated with the JKR model and assuming constant yield stress (see Fig. 11.5). In this case the load is almost entirely due to an external force, while the weight only adds a gravitational force of $4\pi/3 \cdot R^3 \rho g \approx 0.0019$ nN. Negative values of the load indicate the presence of adhesion. Even if we pull on the microsphere, it remains in contact due to attractive forces. Only when pulling with a force stronger than the adhesion force of 850 nN does the particle detach from the surface. The experiment was done with the colloidal probe technique [467]. Please note that for this single microcontact, Amontons' law is not fulfilled! The friction force does not increase linearly with load. Monomolecular layers of hydrocarbons drastically reduce adhesion (to 150 nN) and friction.

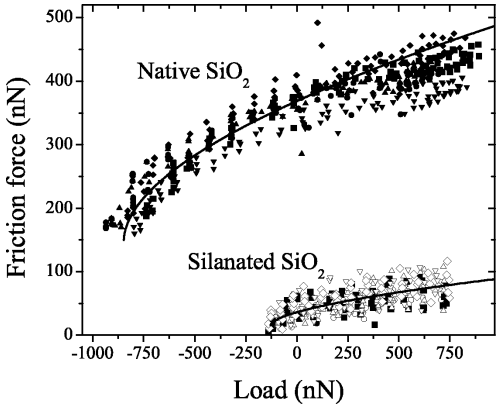



**Figure 11.5:** Dependence of friction on load for a single microcontact. The friction force between a silica sphere of 5 $\mu$m diameter and an oxidized silicon wafer is shown (filled symbols). Different symbols correspond to different silica particles. The solid line is a fitted friction force using a constant shear strength and the JKR model to calculate the true contact area (based on Eq. (6.68)). Results obtained with five different silanized particles (using hexamethylsililazane) on silanized silica are shown as open symbols. Redrawn after Ref. [467].

### 11.1.6 Experimental Aspects

### 11.1.7 Techniques to measure friction

Classical, macroscopic devices to measure friction forces under well-defined loads are called tribometers. To determine the dynamic friction coefficient, the most direct experiment is to slide one surface over the other using a defined load and measure the required drag force. Static friction coefficients can be measured by inclined plane tribometers, where the inclination angle of a plane is increased until a block on top of it starts to slide. There are numerous types of tribometers. One of the most common configurations is the pin-on-disk tribometer (Fig. 11.6). In the pin-on-disk tribometer, friction is measured between a pin and a rotating disk. The end of the pin can be flat or spherical. The load on the pin is controlled. The pin is mounted on a stiff lever and the friction force is determined by measuring the deflection of the lever. Wear coefficients can be calculated from the volume of material lost from the pin during the experiment.

In 1987 Mate et al. [468] used, for the first time, an atomic force microscope (AFM) to measure friction forces on the nanometer scale (review: Ref. [469]). This technique became known as friction force microscopy (FFM) or lateral force microscopy (LFM). To measure friction forces with the AFM, the fast scan direction of the sample is chosen perpendicular to the direction of the cantilever. Friction between the tip and the sample causes the flexible cantilever to twist (Fig. 11.7). This torsion of the cantilever is measured by using a reflected beam of light and a position-sensitive detector in the form of a quadrant arrangement of photodiodes. This new method made it possible for the first time to study friction and lubrication on the nanometer scale.

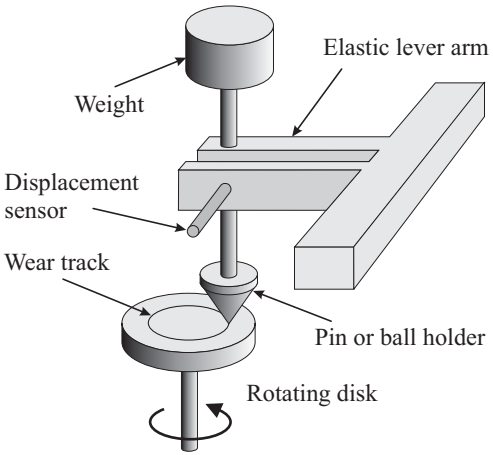



**Figure 11.6:** Schematic of a pin-on-disk tribometer.

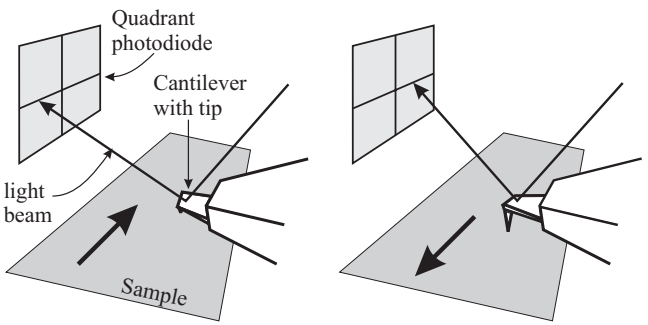



**Figure 11.7:** Working principle of the lateral or friction force microscope (LFM or FFM).

In 1988 a modified surface forces apparatus (SFA) was introduced [470, 471] to analyze friction. The principle of operation of the SFA has already been introduced in Section 6.4. The modified version allowed a relative shearing of the two mica surfaces. In the SFA, the substrate has to have an atomically flat, transparent surface. In most cases mica is used to fulfill these requirements. Although there is a strong limitation in the choice of materials, due to the high resolution in the vertical direction, the SFA has become an important tool to study the friction and lubrication properties of molecularly thin films.

Another tool used to study friction on the molecular scale is the quartz crystal microbalance (QCM) introduced in Section 9.4.1. The QCM has been used to monitor the adsorption of thin films on surfaces via the induced frequency shift [385]. In the years since 1986, Krim and coworkers could show that the slippage of adsorbed layers on the QCM leads to a damping of the oscillator [472]. This damping is reflected as a decrease in the quality factor $Q$ of the oscillator. From the change in $Q$, a characteristic time constant $\tau_s$, the so-called slip-time, can be derived. This corresponds to the time for the moving object's speed to fall to $1/e$, i.e.

a long slip-time stands for low friction. Typical values for $\tau_s$ are in the range of $10^{-9}$s. The limitation of this method is that it works only for weakly adsorbed layers.

### 11.1.8 Macroscopic friction

In his original studies Amontons found a coefficient of friction of 0.3. Meanwhile it has become clear that friction coefficients can assume a whole range of values. With metals, a clear difference exists between clean metal surfaces, oxidized metal surfaces, and metal surfaces with adsorbed gas. Clean metals have coefficients of friction of 3–7. With oxidation, the value decreases to 0.6–1.0. A consequence is that the coefficient of friction can depend on the load. For small loads, friction is determined by the oxide coating. At high loads the microcontacts penetrate the oxide coating, the bare metals come into contact, and the coefficient of friction increases.

There is a similar effect with diamond. In air we have $\mu = 0.1$. In a vacuum, after heating to remove contamination and an oxide layer, $\mu$ rises approximately seven times. This can be explained again with an oxide coating. Many materials, in particular relatively isotropic, soft crystals such as krypton or sodium chloride, lie in the range 0.5–1.0.

Values from tables of friction coefficients always have to be used with caution, since the experimental results not only depend on the materials but also on surface preparation, which is often not well characterized. In the case of plastic deformation, the static coefficient of friction may depend on contact time. Creeping motion due to thermally activated processes leads to an increase in the true contact area and hence the friction coefficient with time. This can often be described by a logarithmic time dependence

$$\Delta A_{real} \propto \ln(1 + t/\tau) \tag{11.8}$$

with a time constant $\tau$.

Very small coefficients of friction are observed on ice. A typical value is $\mu \approx 0.03$. There are several attempts to explain this. Reynolds assumed that the local pressure melts the ice and the water acts as a lubricant. This explanation, still found in some textbooks, is not sufficient. One would then expect that, either the friction does not depend at all on the speed, or that it rises with increasing speed, since there is not sufficient time for melting. The reverse is observed, for example, for skis: at small sliding velocities $\mu = 0.4$ while $\mu = 0.04$ at large velocities. Alternatively, we might assume that local heating melts the ice and creates a water layer. This hypothesis does not agree with results of experiments with differently waxed skis. To fully understand this phenomenon, surface melting well below the bulk melting point has to be taken into account [473].

### 11.1.9 Microscopic friction

In the previous section, we learned that the important processes of friction occur at microcontacts. Thus, to get a better understanding of friction phenomena, one should study friction at the micro- and nanoscale. This field of micro- and nanotribology evolved with the availability of suitable experimental techniques namely FFM, SFA, and QCM (reviews are Refs. [474,475]). The first measurement of friction with atomic resolution using friction force

microscopy was done by Mate et al. [468]. They used a tungsten tip on highly oriented pyrolytic graphite (HOPG) and observed an atomic stick-slip friction, where the periodicity of the stick-slip agreed with the lattice spacing of the graphite. In the following years, atomic stick-slip friction was observed for many other types of crystal surfaces, e.g. NaF, NaCl, AgBr [476], $MoS_2$ [477], stearic acid crystals [478] KBr [479], $CuP_2$ [480]. In these experiments, with moderate load on the AFM tip, no wear of the atomically resolved surfaces was observed.

The possibility of wearless friction had already been postulated by Tomlinson in 1929 [458]. He suggested a simple model that describes the interaction between two surfaces in relative motion (Fig. 11.8). The lower surface (2) is simply represented by a periodic potential $V(x)$. The surface atom (A) is elastically coupled to the upper surface (1) and moves through the potential $V(x)$ as surface (1) moves from the left to the right. At the beginning (A) is pinned at a minimum of the potential (a). As surface (1) moves on, the force on (A) increases (b) until the force gets too high and the atom jumps to the next potential minimum (c). This is an extremely fast step and energy can be dissipated. Energy dissipates via lattice vibrations of the upper body and finally via the generation of phonons. In Tomlinson's model, this dissipation is described by a simple damping term that is proportional to the speed $v_A$ of the atom. This model explains Coulomb's law of friction: The fast, dissipative movement of the surface atom will be largely independent of the relative the speed of the surfaces as long as they move with a speed that is much smaller than this fast relaxation process. In spite of its simplicity, the Tomlinson model has been quite successful for the quantitative interpretation of friction force microscope experiments when extended to two dimensions.

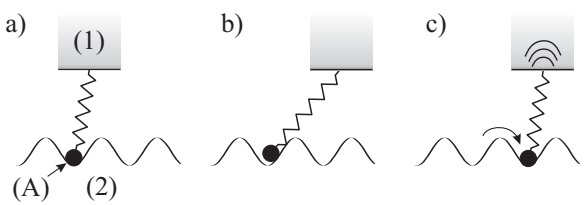



**Figure 11.8:** Illustration of Tomlinson's model.

■ **Example 11.2.** In Fig. 11.9 the experimental results from a friction force microscope experiment is compared to simulations based on an extended two-dimensional Tomlinson model [481]. The tip was assumed to be connected elastically to the holder (coordinates $(x_0, y_0)$ that is scanned with the velocity $v$ relative to the sample surface. The path $(x(t), y(t))$ of the tip was calculated using effective masses $m_x$, $m_y$, spring constants $K_x$, $K_y$, and damping constants $\gamma_x$, $\gamma_y$. The equation of motion for this system is:

$$m_x\ddot{x} = K_x(x_0 - x) - \frac{\partial V(x,y)}{\partial x} - \gamma_x\dot{x}$$
$$m_y\ddot{y} = K_y(y_0 - y) - \frac{\partial V(x,y)}{\partial y} - \gamma_y\dot{y}$$

As interaction potential the authors chose:

$$V(x,y) = -V_0 \left[ 2 \cdot \cos\left(\frac{2\pi}{a}\, x\right) \cos\left(\frac{2\pi}{a\sqrt{3}}\, y\right) + \cos\left(\frac{2\pi}{a\sqrt{3}}\, y\right) \right]$$

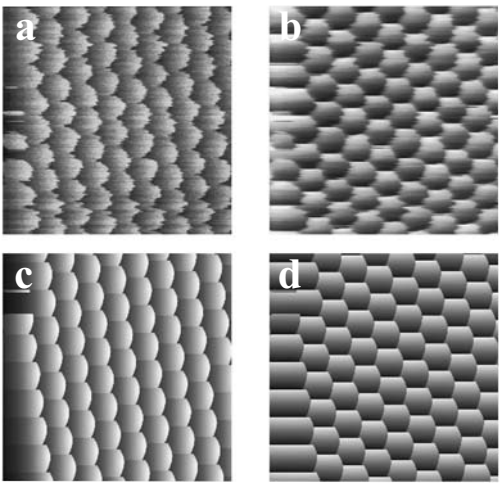



**Figure 11.9:** Friction force microscope pictures (a, b) of a graphite(0001) surface as obtained experimentally with FFM and results of simulations (c, d) of the stick-slip friction using a two-dimensional equivalent of the Tomlinson model. The friction force parallel to the scan direction (a, c) and the lateral force perpendicular to the scan direction (b, d) are shown. The scan size is 20 Å× 20 Å. Pictures taken from Ref. [481] with kind permission from R. Wiesendanger.

In the Tomlinson model for atomic friction, energy dissipation was described by a simple damping term. This phenomenological approach does not reveal anything about the underlying mechanism. For insulating surfaces, phonons are expected to be the relevant dissipative pathway. In the case of conducting surfaces, electronic excitations (e.g. generation of electron–hole pairs) can contribute. The relative contributions of these two mechanisms are still under debate. Measurements with the quartz microbalance on the slippage of nitrogen on a superconductor (lead) showed that, below the transition temperature $T_c$, sliding friction dropped to about half the value compared to temperatures above $T_c$ [472]. If we assume that this drop is due to a complete loss of electronic friction, then electronic and phononic contributions are of the same order of magnitude, at least in this system. Other energy loss mechanisms such as chemical reactions, triboluminescence, or emission of charged particles can also occur [482].

■ **Example 11.3.** The first direct measurement of the energy dissipation through single atoms was achieved by Giessibl et al. [483]. They used an extended version of friction force microscopy, where the tip is oscillated laterally. Dissipation energies were obtained from the frequency shift of the oscillation while measuring the tunneling current between

tip and surface. With this method they were able to measure the energy dissipation during interaction of the tip with single adatoms of a Si(111)7 $\times$ 7 surface. When the tip was placed above an adatom, the dissipation was low. A positive frequency shift of the oscillation indicated an increase of the effective spring constant of the oscillator. This was interpreted as the formation of a "bond" between tip and adatom. When the tip was positioned further away from the adatom the "bond" was formed and destroyed in each oscillation cycle and more energy was dissipated. In this case the large deflection of the adatom from its equilibrium position leads to strong vibrations and generation of phonons.

The validity of Coulomb's law has been verified also on the nanoscale: Zwörner et al. [484] showed that, for different carbon compound surfaces, friction does not depend on sliding velocity in the range between 0.1 $\mu$m/s and up to 24 $\mu$m/s. At low speeds, a weak (logarithmic) dependence of friction on speed was observed by Gnecco et al. [485] on a NaCl(100) surface and by Bennewitz et al. [486] on a Cu (111) surface. This can be modeled when taking into account thermal activation of the irreversible jumps in atomic stick-slip [487].

For macroscopic bodies, the real and apparent contact areas are different. It was the deformation of surface asperities that accounted for Amontons' law. If considering a single nanocontact, apparent and real contact areas are identical. Assuming that the shear strength of the contact does not depend on the normal load, it is possible to calculate the load dependence of the friction force by applying contact mechanics to predict the true contact area. In general this will lead to a non-linear dependence of $F_F$ on $F_L$. Experiments on the load dependence of friction for single nanocontacts have been carried out by several authors using friction force microscopy. The resulting friction force for different systems could be described using continuum theories such as the JKR model (Eq. (6.68)) to calculate the true contact area [488] [489, 490]. This result was confirmed by determining independently the true area of contact from the electric conductance of the contact [491].

In experiments with friction force microscopy, the tip forms a contact of a few nanometers in diameter with the substrate, a so-called *nano*contact. In reality, friction of macroscopic bodies is determined by the interaction via *micro*contacts. One possibility of extending the method of friction force microscopy to larger contact areas is the use of the colloidal probe technique, where a small sphere is attached to the end of an atomic force microscope cantilever (see Section 6.4). Even for microcontacts, the proportionality between the true area of contact and the friction force was observed (see example 11.1).

■ **Example 11.4.** McGuiggan et al. [492] measured the friction on mica surfaces coated with thin films of either perfluoropolyether (PFPE) or polydimethylsiloxane (PDMS) using three different methods: The surface forces apparatus (radius of curvature of the contacting bodies $R \approx 1$ cm); friction force microscopy with a sharp AFM tip ($R \approx 20$ nm); and friction force microscopy with a colloidal probe ($R \approx 15$ $\mu$m). In the surface force apparatus, friction coefficients of the two materials differed by a factor of $\approx 100$ whereas for the AFM silicon nitride tip, the friction coefficient for both materials was the same. When the colloidal probe technique was used, the friction coefficients differed by a factor of $\approx 4$. This can be explained by the fact that, in friction force experiments, the contact pressures are much higher. This leads to a complete penetration of the AFM tip through the lubrication layer, rendering the lubricants ineffective. In the case of the colloidal probe the contact pressure is reduced and the lubrication layer cannot be displaced completely.

## 11.2 Lubrication

The reduction of friction by lubricants was a prerequisite for the industrial revolution. Lubrication helps to reduce energy consumption and increase the lifetime of machines by minimizing wear. Without lubricants, almost no machine made of metal would work. It is not surprising that the phenomenon of friction and lubrication was of interest since ancient times. We know that the Egyptians wetted the sand on which they transported their stones, to reduce friction [493].

Depending on the thickness of the lubricating layer, we distinguish between two different lubrication regimes. In hydrodynamic lubrication the lubrication layer is thicker than the maximum height of the surface asperities resulting in a complete separation of the friction partners. In boundary lubrication the lubrication layer is typically only a few molecular layers thick and therefore thinner than the surface roughness. In many practical applications we are between the two extremes, which is referred to as mixed lubrication.

### 11.2.1 Hydrodynamic lubrication

The principle of hydrodynamic lubrication can be easily understood from a simple configuration as shown in Fig. 11.10, where a slider is moving with constant velocity $v$ over a substrate in the presence of a lubricant. The viscous drag force on the lubricant film between the surfaces causes the formation of a hydrodynamic wedge. In such a system, the coefficient of friction depends on the fluid dynamics, in particular on the viscosity $\eta$ of the lubricant. For this reason, hydrodynamic lubrication is also called fluid lubrication. A well known example for hydrodynamic lubrication is the effect of aquaplaning. When a car is driven at high speed on a wet road, the water forced between the surfaces of the tyres cannot escape and separates the car from the road; traction is lost. The build up of the lubrication film can either be solely due to the relative movement of the bearing surfaces or can be achieved by active pumping of the lubricant.

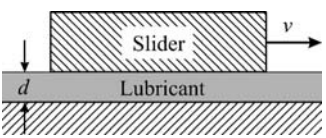



**Figure 11.10:** Two solids with parallel planar surface sliding over each other across a lubricating film.

Since in hydrodynamic lubrication the friction force is completely determined by the viscous friction of the lubricant, the coefficient of friction can be calculated from hydrodynamics using the Navier–Stokes equations. This had already been done in 1886 when Reynolds published his classical theory of hydrodynamic lubrication [494]. The friction force $F_F$ between two parallel plates of area $A$ separated by the distance $d$ is given by:

$$F_F = \frac{A}{d} \cdot \eta \cdot v \tag{11.9}$$

where $v$ is the relative speed and $\eta$ is the viscosity of the lubricant. This linear dependence on velocity can be derived from the Navier–Stokes equations using two assumptions: First, the flow of the lubricant is laminar and not turbulent. This is fulfilled in most lubrication situations

because the Reynolds number (Eq. (7.35)) $Re = \rho v d/\eta$ is usually much lower than unity due to the small distances $d$. Typically lubricant layers are of the order of 1 $\mu$m thick. Second, the lubricant is a Newtonian fluid, i.e., the viscosity does not depend on the shear rate. The shear rate $\dot{\gamma}$ for two parallel plates is the velocity divided by the distance: $\dot{\gamma} = v/d$. This is fulfilled for most lubricants except at extremely high shear rates. At very high shear rates, the viscosity might decrease, a phenomenon known as shear thinning.

■ **Example 11.5.** A smooth quadratic steel plate of 10 cm length slides on top of another steel plate with a speed of 1 m/s. Both plates are separated by a lubricating film of hexadecane (viscosity at 25°C: 3.03 mPa s) of 4 $\mu$m thickness. The force required to maintain sliding is

$$F_F = \frac{(0.1\ \mathrm{m})^2 \cdot 3.03 \times 10^{-3}\mathrm{sPa} \cdot 1\ \mathrm{ms}^{-1}}{4 \times 10^{-6}\mathrm{m}} = 7.58\ \mathrm{N}$$

For a fixed geometry, the friction force depends solely on the viscosity of the lubricant. We could try to decrease the viscosity of the lubricant to reduce friction. There is, however, a limit to this: The lubrication film thickness must always be kept higher than the surfaces asperities. Otherwise the surfaces will come into direct contact, resulting in much higher friction. Therefore, it is common to use an oil with a viscosity that is just high enough to maintain a continuous lubrication layer.

In practical applications, the increase of viscous friction with speed is often lower than expected from Eq. (11.9). The explanation is that friction leads to an increased temperature of the lubricant which reduces the viscosity. For most lubricants the temperature dependence of the viscosity is given by

$$\eta = \eta_0 \cdot e^{E_a/k_B T} \tag{11.10}$$

Here, $E_a$ is the effective activation energy. This leads to an inherent stability of hydrodynamic lubrication, as a thinning of the lubricant at higher temperatures reduces the friction and therefore avoids further heating.

In gears and ball bearings, extremely high local pressures occur, at least for short periods of time for which the lubricant has no time to flow away from such a high-pressure region. Under these conditions, the above theory of hydrodynamic lubrication would predict a lubrication layer thickness that is smaller than the surface corrugation. However, experiments show that fluid film lubrication still holds under such conditions. To understand this phenomenon, we have to take two effects into account, which lead to elasto-hydrodynamic lubrication:

- The pressure dependence of the viscosity. The viscosity of most lubricants increases roughly exponentially with increasing pressure:

  $$\eta = \eta_0 e^{\alpha P} \tag{11.11}$$

  Here, $\alpha$ is a constant and $\eta_0$ is the viscosity at zero pressure. Eq. (11.11) tells us that the higher the pressure becomes, the harder it becomes to squeeze the lubricant out of the gap. In addition, there may even be a phase change of the lubricant that leads to solidification.

- The solids confining the lubricant are never perfectly stiff, but deform elastically at high pressure. As a result the area over which hydrodynamic pressure is spread, increases. This effect tends to prevent direct contact between the surfaces.

In fact, increasing the load under these conditions may have little effect on the thickness of the lubricating film. The oil film becomes actually more rigid than the metal surfaces. Therefore, an increase in load does not much decrease the film thickness but increases the pressurized area by deformation of the bearing surfaces. In practice, this picture is complicated by possible high local temperatures and high shear rates. Under these conditions, the lubricant may no longer behave like a Newtonian fluid and effects like shear thinning may become important.

The great advantage of (elasto-)hydrodynamic lubrication is that friction can be quite low and in principle there is no wear of the moving parts. Furthermore, the underlying mechanism is well understood and allows the calculation of the necessary parameters for the construction of bearings. In practice, however, wear may still occur during starting / stopping, where the hydrodynamic lubrication breaks down. In the case of elasto-hydrodynamic lubrication, the repeated deformation of the bearing surfaces may lead to fatigue failure.

### 11.2.2 Boundary lubrication

At low sliding velocities and high loads, the lubricating film is squeezed out of the gap. This leads to so-called boundary lubrication. Friction coefficients under these conditions are typically 100 times higher than under hydrodynamic lubrication conditions, but still substantially smaller than for dry friction under UHV conditions. This is due to the fact that the surfaces are still wetted by molecular layers of the lubricant, even under conditions where the local stress is high enough to deform the surface asperities. Under these conditions friction depends more on the chemical constitution of the lubrication layer than on its viscosity.

One effect of a lubricant is to reduce adhesion between the solids. Adhesion between solids is usually dominated by van der Waals forces. Hydrocarbons have a small Hamaker constant and their presence leads to a reduction in the adhesion and hence friction. Films of only a monomolecular thickness are sufficient to have a pronounced effect. This shows up when we measure friction between solids, which are coated with monomolecular layers (see example 11.1). In that case, friction can be as small as friction with plenty of lubricant. A monomolecular film affects significantly the frictional properties [495]. At least with metals it can be shown that the number of microcontacts is not changed by the lubricant. Only the contact intensity is reduced. The reduced van der Waals attraction can thereby diminish the actual contact area.

It should be noted, that most “clean” surfaces (unless prepared and kept in UHV) are covered by a thin layer of contamination. As a consequence, in experiments performed under ambient conditions, we have to consider that boundary lubrication rather than dry friction might be present.

In practical applications we often encounter a combination of boundary and hydrodynamic lubrication which is called mixed lubrication. For example, bearings that are usually lubricated hydrodynamically, experience mixed lubrication when starting and stopping. This is shown in the Stribeck diagram (Fig. 11.12): At low speeds, boundary lubrication with high friction

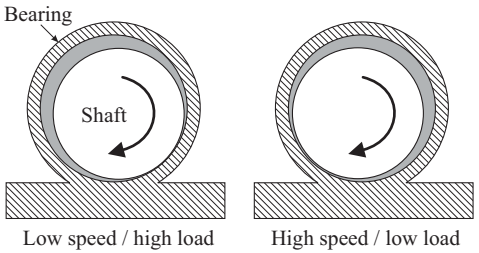



**Figure 11.11:** Different lubrication situations in a journal bearing. Left: At low velocities and high loads, boundary lubrication with a high friction coefficient dominates. The shaft climbs the journal on the right side. Right: At high speeds and low loads hydrodynamic lubrication leads to much lower friction. The build up of the hydrodynamic wedge moves the shaft to the upper left.

dominates. With increasing speed a hydrodynamic lubricant film is created that significantly reduces the friction. The linear increase of friction at even higher speeds is due to the internal friction (viscosity) of the lubricant, according to Eq. (11.9).

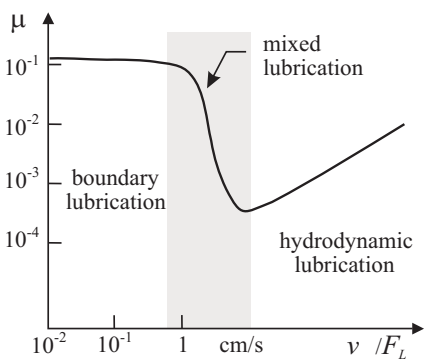



**Figure 11.12:** Stribeck diagram: Plot of friction coefficient $\mu$ versus $\eta_k v/P$,where $\eta_k$ is the kinematic viscosity, $v$ the velocity, and $P$ the contact pressure. From left to right there are three different friction regimes: Boundary lubrication with high friction and wear, mixed lubrication with medium friction and wear, and hydrodynamic lubrication with low friction and wear.

In contrast to the self-stability of hydrodynamic lubrication, the heating effects in mixed lubrication tend to destabilize the film. An increase in temperature reduces the viscosity, which in turn reduces the film thickness, leading to increased probability of intermittent contact. This further increases friction and the temperature rises even more.

### 11.2.3 Thin film lubrication

The design of precision components with ultrasmooth surfaces, for example in the field of micro-electromechanical systems (MEMS) and nanotechnology, boosts the use of lubricating

films of molecular thicknesses. Under these conditions, the validity of continuum theories to describe the hydrodynamics of the lubricant, is questionable. This is a new lubrication regime, denoted as thin film lubrication (review: Ref. [496]).

As we have seen in Section 6.6.1 such confined liquids may behave quite differently from the bulk lubricant. Near the surfaces, the formation of layered structures can lead to an oscillatory density profile (see Fig. 6.12). When these layered structures start to overlap, the confined liquid may undergo a phase transition to a crystalline or glassy state, as observed in surface force apparatus experiments [471, 497–500]. This is correlated with a strong increase in viscosity. Shearing of such solidified films, may lead to stick-slip motions. When a critical shear strength is exceeded, the film liquefies. The system relaxes by relative movement of the surfaces and the lubricant solidifies again.

In thin lubricating films we have to take another phenomenon into account: slippage. In fluid mechanics we usually rely on the assumption that when liquid flows over a solid surface, the liquid molecules adjacent to the solid are stationary relative to the solid. This is the so-called no-slip boundary condition. Using the no-slip boundary condition the flow of macroscopic systems can successfully be described. For liquids with low liquid–solid interaction computer simulations [501–503] and experiments [504–507] showed that this condition is not necessarily fulfilled and we might, in fact, observe slippage. At first glance this case is not relevant with respect to lubrication because lubricants have to wet the solid surfaces, which implies that the liquid–solid interaction is strong compared to liquid–liquid interactions. Recently, however, slip was observed even for strong liquid–solid interactions at high shear rates [508, 509]. Thus, slip might in fact occur and reduce friction under conditions of thin film lubrication.

### 11.2.4 Lubricants

The main purpose of lubricants it to reduce friction and minimize wear. In 1999 about 37 million tons of lubricants were consumed, worldwide. Most lubricants consist of a base oil and additives that are used to improve the desirable properties. The two main sources of base oils are the refining of petroleum crude oil and the synthesis of lubricants from lower molecular weight materials. Refined base oils mainly consist of paraffins and naphthenes. Oils from harvestable natural raw material gain increasing acceptance because of their environmental compatibility. Another type of lubricant is grease. Greases are dispersions of a thickening agent — usually a soap — in a liquid lubricant. Under special conditions (very high pressures, high temperatures, UHV) solid lubricants such as $MoS_2$ or graphite have to be used. An extensive overview on the topic of lubricants is given by Ref. [510].

Lubricants are characterized by several parameters. The most important one is the **viscosity**, sometimes also called dynamic viscosity. In lubrication industry, often the kinematic viscosity is used. The kinematic viscosity is the dynamic viscosity divided by the density of the liquid: $\eta_k = \eta/\rho$. It is given in units of $m^2/s$ or centistoke (1 centistoke = 1 $mm^2/s$ = 0.001 $m^2/s = 10^{-2}$stoke).

As already mentioned, the viscosity of a base oil decreases with increasing temperature. Therefore, it is important to know, not only the viscosity at a certain temperature, but also how much it changes within a temperature range given by operating conditions. To characterize the temperature dependence of viscosity in 1929 the American Society for Testing and

Materials (ASTM) introduced the so-called **viscosity index** (VI). It is based on the values of kinematic viscosity at 40°C and 100°C, which are compared to the respective viscosities of two reference oils. One reference oil is standard paraffin oil. Its viscosity only weakly depends on temperature. It was given a viscosity index of 100. The other reference is a standard naphthenic oil with a high temperature dependence. Its viscosity index is defined as zero. A low viscosity index indicates a relatively strong dependence of viscosity on temperature, a high viscosity index, a small dependence.

**Volatility** is a second important property of a lubricant. It should be as low as possible. Volatility is determined by the base oil because it makes up for most of the material, and cannot significantly be influenced by the additives. Therefore, base oils are usually low volatile oils with high boiling points.

For applications, the **aging behavior** is important. How long does a lubricant maintain its properties? This will strongly depend on the operating conditions, the temperature, oxidative conditions, contamination with water, acidic combustion residues or particulate matter. Apart from contaminations, there are two main aging mechanisms: oxidation and thermal decomposition. Oxidation of hydrocarbons generates high molecular compounds [511]. The viscosity of the oil increases and finally insoluble polymers form sludge and varnish-like deposits. Oxidation of the base oil is, in most cases the most critical aging process. Thermal decomposition (cracking) in the absence of oxygen, splits hydrocarbons into smaller molecules.

The performance of lubricants can be significantly improved by the use of additives, which can make up from ppm to 30%, (w/w) of the lubricant. Typical values of total additive content are in the order of 5% . Often different additives are combined to improve several properties. Choosing the right combination can have synergistic effects on the overall performance, but a wrong choice can render some of them inactive. An important issue with additives can be the compatibility of the additives with the solid surfaces or with polymer or rubber seals. In table 11.1 the most important types of additives are listed. The list also illustrates the main practical problems encountered when making lubricants.

## 11.3 Wear

Wear is defined as the progressive loss of material from a body caused by contact and relative movement of a contacting solid, liquid, or gas. The importance of understanding and minimizing wear in technical designs is obvious, but still today there are no reliable methods to theoretically predict the lifetime of a new design. Several equations are used to describe wear rates. One example is Archard's well-known law of adhesive wear [512], which describes the material loss per time:

$$\Psi = k_w \cdot \frac{F_L v}{H} \tag{11.12}$$

Here, $H$ the hardness of the material and $k_w$ the wear coefficient (in $m^3/s$). The practical usefulness of such equations, however, is limited. Ludema [452] comments on the results of an extensive literature scan [513], that yielded more that 300 different equations describing wear under different conditions: "Many of the equations appeared to contradict each other and very few equations incorporated the same array of variables. It is common to find, for example,

**Table 11.1:** Types of lubrication additives.

| | |
|---|---|
| Antioxidants | Prevent oxidation by neutralizing radicals. |
| Dispersants | Keep contaminants of the lubricant in suspension and avoid their aggregation. Dispersants are amphiphilic molecules. Their long hydrocarbon tail helps to solubilize polar molecules in the base oil. The polar head group interacts with contaminants and facilitates the formation of (inverse) micelles around them. |
| Viscosity modifiers | Increase the viscosity or viscosity index of the lubricant. Typically polymers are used. |
| Pour point depressants | The crystallization of paraffin wax in the base oil can lead to a gelation of the lubricant at low temperature. Pour point depressants cannot prevent crystallization but change the shape of the crystals from a needle-like to a densely packed, rounded one. This leads to a much better flow behavior. |
| Antifoam agents | Prevent foam formation. Foaming of lubricants is detrimental for two reasons. First, in an oil-circulation system foaming may lead to a break-down of the lubrication transport. Second, the increased surface area in a foam accelerates oxidation. |
| Demulsifiers | Without demulsifiers water contaminations will form water-in-oil emulsions. Today mainly poly(ethylene glycols) are used. |
| Friction modifiers | Compounds like fatty acids form physisorbed layers on the metal surfaces. They reduce friction under conditions of mixed lubrication and help to avoid stick-slip motion. |
| Antiwear (AW) and extreme pressure (EP) additives | AW and EP additives can be chemisorbed to metal surfaces and help to avoid direct contact of the metal. |
| Corrosion inhibitors | Corrosion inhibitors cover metal surfaces and thereby isolate them from attack by oxygen and moisture. Usually corrosion inhibitors are amphiphilic molecules that bind with their polar group to the metal. |

Young's Modulus in the numerator of some equations and in the denominator of others". In addition, the correlation between friction and wear is far from being clear. As already mentioned, atomic friction can occur without any wear. During running-in of machines, mild wear may lead to a reduction in friction. In some situations with high wear, friction may be relatively low.

The terminology used in the description of wear is diverse. Sometimes terms are used that describe the damage symptoms, in other cases they result from the attempt to describe the complex sequence of underlying events in a few words. The following list is certainly not complete and concentrates on the most common terms.

**Abrasion** occurs when one material is in contact with a harder material. Surface asperities of the harder material cut, plough, or indent characteristic scratches or grooves into the softer material (two-body abrasion). Abrasion can also be caused by hard particles that are trapped in between two surfaces (three-body abrasion). Irregular patterns of small indentations are formed. Contamination in the lubricants can significantly contribute to this type of abrasion.

**Erosion** is caused by the impingement of a fluid or air stream containing solid particles. The eroding effect of a sandblast unit is a well-known example. The underlying mechanism may vary, depending upon the the ratio of particle to target hardness [514]. On the surface of metals, the impingement of the particles leads to work hardening of the surface and subsequently to lateral cracking or small-scale chipping [515].

Excessive load, reduced fluid viscosity, and low speed can cause a breakdown of the lubrication layer to the point of a metal–metal contact. Direct **adhesive contact** can lead to molecular interaction strong enough to tear material out of the surface. This process is also known as "cold welding". Metal surfaces under ambient conditions are usually covered with oxide layers that prevent cold welding at moderate loads.

**Surface fatigue** is usually the result of high periodic loads in the contact zone that lead to changes of the material structure and finally to material failure at the surface.

**Fretting wear** is observed in systems where two surfaces perform an oscillatory relative sliding with small amplitude (= fretting). This is often due to vibrations. In principle the same wear behavior as for normal sliding, should occur. However, wear particles tend to remain within the immediate contact region where they can act as an abrasive and lead to increased wear.

Apart from oxidation of the lubricant and the metal surfaces, there can be complex **tribochemical reactions**. Chemical reactions at the surfaces can be stimulated by different factors. One factor is heating due to friction. This can either be a global effect (elevated mean temperature of surfaces and lubricant) or a localized phenomenon. Especially in situations where mixed or boundary lubrication exists, the direct contact of surface asperities can lead to high flash temperatures. At these "hot spots" temperatures in excess of 1000°C promote chemical reactions and surface melting. Other factors promoting chemical reactions are:

- Triboemission of low energy electrons. This leads to formation of highly reactive radicals [516].
- Removal of protective oxide layers.
- Increase of the reactive surface by roughening.
- Creation of "dangling bonds" on the surface by plastic deformation.

Imploding gas or vapor bubbles can damage the material surfaces. **Cavitation** often intensifies corrosive processes and can lead to rapid degradation of the surfaces. Examples are ship propellers or the vibrating cylinder walls of engines. In lubricated systems cavitation can be minimized by using high boiling point oils, avoiding the drag-in of air, and the use of surface active additives to reduce the size of possible air bubbles.

## 11.4 Summary

- Amontons' law of macroscopic, dry friction states that the friction force is proportional to the load and does not depend on the apparent contact area:

  $$F_F = \mu \cdot F_L$$

  The proportionality constant $\mu$ is the friction coefficient. A microscopic analysis shows that the friction force is, in fact, proportional to the true contact area that, in turn, is proportional to the load. The friction coefficient is in general higher for static than for dynamic friction.

- The second empirical law for dry, macroscopic friction is that of Coulomb: Friction does not depend on the sliding velocity.

- The availability of new experimental methods at the end of the 1980's allowed us to study friction on the atomic scale and created the new field of nanotribology. The observed wearless friction on this scale can be understood using the model of Tomlinson where the "plucking action of one atom on to the other" leads to energy dissipation via the generation of phonons.

- Lubrication is used to reduce friction and wear. Depending on the thickness of the lubrication layer, different regimes are distinguished. Hydrodynamic lubrication prevails if the lubrication layer is thick enough to completely separate the surfaces. In this case friction is determined by the viscosity of the lubricant. If the surfaces are only separated by adsorbed or oxide layers, boundary lubrication is acting. In systems where the thickness of the lubrication layer is of the order of the surfaces roughness and intermittent contact of the surfaces occurs, one speaks of mixed lubrication. With the evolution of nanotechnology, thin film lubrication is gaining importance, where the lubrication layer thickness approaches molecular dimensions. Lowest friction coefficients can be achieved in the case of rolling friction as, for example, in ball bearings.

| Operating conditions | friction coefficient |
|---|---|
| Dry friction | 0.1 $\ldots \geq 1$ |
| Boundary lubrication | 0.01 $\ldots$ 0.2 |
| Mixed lubrication | 0.01 $\ldots$ 0.1 |
| Hydrodynamic lubrication | 0.001 $\ldots$ 0.01 |
| Rolling friction | 0.001 $\ldots$ 0.005 |

- Lubricants consist of base oils that may either stem from refining of crude oil or be synthetic and a formulation of additives to achieve the desired performance of the lubricant.
- Wear is the progressive loss of material caused by contact and relative motion. In spite of the technical and economic importance, there is no general theory to describe wear, and engineers still have to rely on their extensive empirical knowledge to construct machines with low wear.

## 11.5 Exercises

1. A block of metal with a mass of 5 kg is put on an plate of the same metal 1 meter from the left end of the plate. The coefficient of static friction of the material is 0.5,the coefficient of kinetic friction is 0.4. The right end of the plate is slowly being lifted to incline the plate until the block begins to slide. At which inclination angle does the block begin to slide? If the plate is kept at this angle, what speed will the block have when it reaches the left end of the plate?
2. A steel cube of 10 cm side length is placed on a steel plate. The yield stress of steel is $\sigma_y \approx 10^9\,\mathrm{N/m^2}$. Estimate the area of real contact. If we assume a typical area of $10\ \mu\mathrm{m}^2$ per microcontact, how many junctions exist between the plate and the block?
3. A cylindrical axis of 1 cm diameter and 10 cm length is rotating in a precision bearing with an inner diameter, which is only 6 $\mu$m larger. We use octadecane as the lubricant. Due to friction it heats up to 50°C, where its viscosity is 2.49 mPa s. The axis rotates with 80 rotations per second. What is the torque we have to apply to overcome the viscous friction?

# 12 Surfactants, micelles, emulsions, and foams

## 12.1 Surfactants

In this Chapter we discuss interfacial phenomena, which are dominated by the properties of soluble amphiphilic molecules also called amphiphiles [517]. Amphiphilic molecules are present in everydays life and they are essential in many industrial processes. Examples of applications are their use in cleaning, detergency, as emulsifiers for creams in the cosmetic and pharmaceutical industry, and as flotation agents in the mining industry.

Soluble amphiphiles are also known as detergents, tensides, or surfactants. Perhaps the most descriptive of these words is the word surfactant, which is a contraction of the phrase "surface active agent". The term soap is usually restricted to the alkali metal salts of long-chain fatty acids (Table 12.1). The term "amphiphile" indicates that one part of the molecule likes a certain solvent while the other part likes another solvent and the two solvents are immiscible. Usually one solvent is water and the water-loving part is called hydrophilic. The other part is hydrophobic. It does not like to be in water and prefers to be in an oily environment or air. The hydrophobic part usually consists of a long, straight alkyl chain ($CH_3(CH_2)_{n_C-1} \sim$, $n_C$ = 8–20). For special applications the hydrocarbons might be completely or partially fluorinated.

Surfactants are divided into four groups (Table 12.1):

- **Anionic** surfactants have a hydrophilic group, which carries a negative charge such as a carboxylate, sulfonate, or sulfate group. An important and widely used anionic surfactant is sodium dodecylsulfate (SDS), $C_{12}H_{25}OSO_3Na$, which belongs to the class of sodium alkylsulfates. In water they dissociate according to

  $$C_{12}H_{25}OSO_3Na \rightarrow C_{12}H_{25}OSO_3^- + Na^+$$

  Another anionic surfactant is sodium dodecanoate, $C_{11}H_{23}CO_2Na$. It belongs to the class of sodium alkylcarboxylates, also called fatty acids or soaps. Just as with SDS in water at neutral pH, the alkali metal dissociates as a cation so that the surfactant becomes negatively charged.

- **Cationic** surfactants bear a positive charge in their hydrophilic part. Examples are dodecyl trimethylammonium bromide, $C_{12}H_{25}N(CH_3)_3Br$, and hexadecyl trimethylammonium bromide, which dissociate in water according to

  $$C_{16}H_{33}N(CH_3)_3Br \rightarrow C_{16}H_{33}N^+(CH_3)_3 + Br^-$$

  with the positive charge localized at the nitrogen. Hexadecyl trimethylammonium bromide is also called cetyl ammonium bromide (CTAB).

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

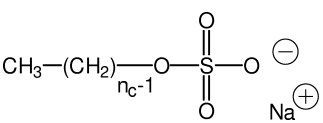
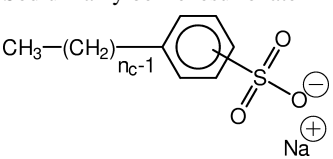
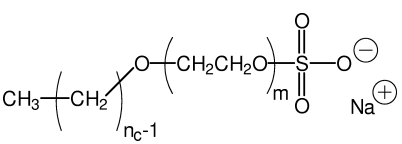
- **Nonionic** surfactants are not charged. The hydrophilic part derives its water solubility from highly polar groups such as polyethylene oxide or sugars. Alkylethylene oxides, also called alkylethylene glycols, belong to the group of nonionic surfactants. Two examples are $C_{10}H_{21}(OCH_2CH_2)_8OH$ and $C_{12}H_{25}(OCH_2CH_2)_6OH$. Alkylethylene glycols are usually written as $C_{n_c}E_{n_e}$, where the index $n_c$ indicates the number of carbon atoms in the alkyl chain and $n_e$ is the number of ethylene oxide units in the hydrophilic head. The short notations for the above molecules are $C_{10}E_8$ and $C_{12}E_6$, respectively.

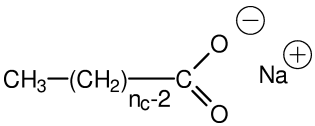
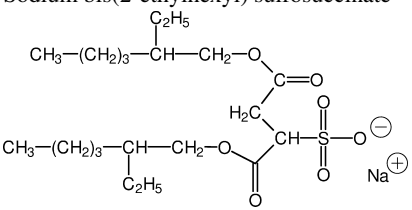
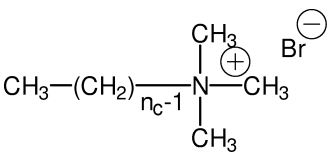
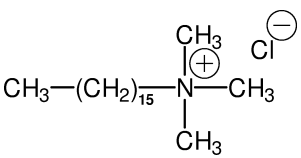
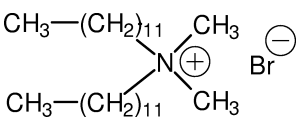
  Another class of nonionic surfactants are the alkyl glycosides, also known as alkyl polyglycosides [518]. They consist of mono- or oligosucrose, glucose, or sorbitol as hydrophilic head group with an alkyl chain as hydrophobic tail.

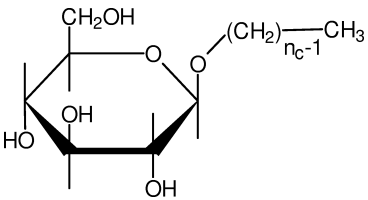
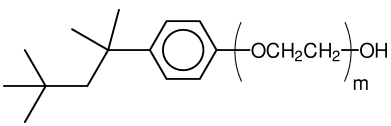
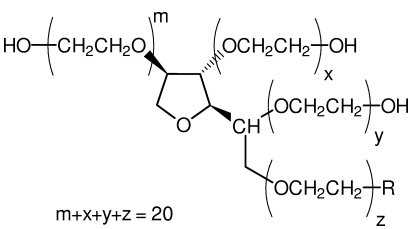
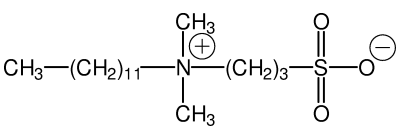
- **Amphoteric** (or **zwitterionic**) surfactants carry a positive and a negative charge so that the net charge is zero. Some lipids, such as phophatidylcholine, are zwitterionic. Since lipids are practically not soluble in water, in a strict sense they do not belong to the class of surfactants.

Most surfactants used are anionic surfactants followed by nonionic ones. Cationic surfactants often pose environmental problems since they are not easily biodegradable. Amphoteric surfactants are expensive and therefore they are only used for special applications.

**Table 12.1**: Structure of common surfactants and critical micelle concentrations (CMCs in mM) in water at 25°C (no added salt). The CMCs were taken from Refs. [109, 519, 520].

| Surfactant | | CMC |
|---|---|---|
| Anionic | | |
| Sodium alkylsulfate  | $n_C = 8$, sodium octylsulfate | 139 |
| | $n_C = 10$, sodium decylsulfate | 34 |
| | $n_C = 11$, sodium undecylsulfate | 17 |
| | $n_C = 12$, sodium dodecylsulfate (SDS) | 8.9 |
| Sodium alkylbenzenesulfonate  | $n_C = 7$, sodium heptylbenzene sulfonate | 24 |
| | $n_C = 8$, sodium octylbenzene sulfonate | 12 |
| | $n_C = 12$, sodium dodecylbenzene sulfonate | 3.6 |
| Sodium alkylether sulfate  | AES $n_C = 12 - 14, m = 2 - 4$ | |

| Surfactant | | CMC |
|---|---|---|
| Sodium alkylcarboxylate  | Sodium salt of | |
| | $n_C = 10$, decanoic acid | 100 |
| | $n_C = 11$, undecanoic acid | 50 |
| | $n_C = 12$, dodecanoic (lauric) acid | 25 |
| | $n_C = 13$, tridecanoic acid | 13 |
| | $n_C = 14$, tetradecanoic (myristic) acid | 6.3 |
| | $n_C = 16$, hexadecanoic (palmitic) acid | 1.8 |
| | $n_C = 18$, octadecanoic (stearic) acid | |
| | $n_C = 20$, eicosanoic (arachidic) acid | |
| | $n_C = 22$, docosanoic (behenic) acid | |
| Sodium bis(2-ethylhexyl) sulfosuccinate  | Aerosol OT (AOT) | 1.4 |
| <u>Cationic</u> | | |
| Alkyltrimethylammonium bromide  | $n_C = 10$, decyl trimethylammonium bromide | 66 |
| | $n_C = 12$, dodecyl trimethylammonium bromide | 15 |
| | $n_C = 14$, tetradecyl trimethylammonium bromide (TTAB) | 3.5 |
| | $n_C = 16$, hexadecyl trimethylammonium bromide (CTAB) | 0.9 |
| Alkyltrimethylammonium chloride  | hexadecyl trimethylammonium chloride | 1.3 |
| Dialkyldimethylammonium bromide  | didodecyldimethylammonium bromide (DDAB) | 0.15 |

| Surfactant | | CMC |
|---|---|---|
| <u>Nonionic</u> | | |
| Alkylethylene glycol | | |
| $C_{10}H_{21}(OCH_2CH_2)_4OH$ | $C_{10}E_4$ | 0.79 |
| $C_{10}H_{21}(OCH_2CH_2)_6OH$ | $C_{10}E_6$ | 0.9 |
| $C_{10}H_{21}(OCH_2CH_2)_8OH$ | $C_{10}E_8$ | 1.0 |
| $C_{12}H_{25}(OCH_2CH_2)_8OH$ | $C_{12}E_8$ | 0.071 |
| $C_{14}H_{29}(OCH_2CH_2)_8OH$ | $C_{14}E_8$ | 0.009 |
| Alkylglucosides | $n_C = 8$, octyl-$\beta$-D-glucoside | 25 |
|  | $n_C = 10$, decyl-$\beta$-D-glucoside | 2.2 |
| | $n_C = 12$, dodecyl-$\beta$-D-glucoside | 0.19 |
| Poly(ethylene oxide) | $m = 7, 8$, Triton$^{(R)}$ X-114 | 0.20 |
| *iso*-octylphenyl ether | $m = 10$, Triton$^{(R)}$ X-100 | 0.24 |
| | $m = 40$, Triton$^{(R)}$ X-405 | 0.81 |
|  | | |
| Poly(ethylene oxide) | R=$\sim$OCO$(CH_2)_{10}CH_3$: monolaurate | 0.08 |
| sorbitan monoalkanoate | Tween$^{(R)}$20 | |
| | R=$\sim$OCO$(CH_2)_{16}CH_3$: monostearate | 0.0027 |
| | Tween$^{(R)}$60 | |
|  | | |
| <u>Zwitterionic</u> | | |
| Alkyldimethylpropanesultaine | N-dodecyl-N,N-dimethyl propanesultaine | |
|  | | |

Beside the conventional surfactants with one polar head group and one nonpolar tail, dimeric and oligomeric surfactants have attracted considerable interest in academia and industry [521]. Dimeric surfactants, also called Gemini surfactants, are made up of two amphiphilic moieties connected closely to the head group by a spacer group (Fig. 12.1). In bolaform surfactants the connection is in the middle of the alkyl chain or close to the end, so that they can be considered as two polar head groups connected by a long hydrophobic chain.

More than two surfactants can be put together to form tri,- tetra- or polymeric surfactants. Trimeric or even tetrameric surfactants show properties often superior to monomeric surfactants. Besides, they are intermediate between conventional surfactants and polymeric surfactants. In a normal polymeric surfactant each monomer unit is amphiphilic. Another type of polymeric surfactant, called block copolymer [522], consists of at least two parts. One part is made of monomer type A, the other part is made of monomer B. If A is polar and B nonpolar, the blockcopolymer will be strongly surface active and show many properties of a conventional surfactant. If there are two different blocks we talk about a diblock copolymer. In the following part of this chapter we concentrate on conventional surfactants.

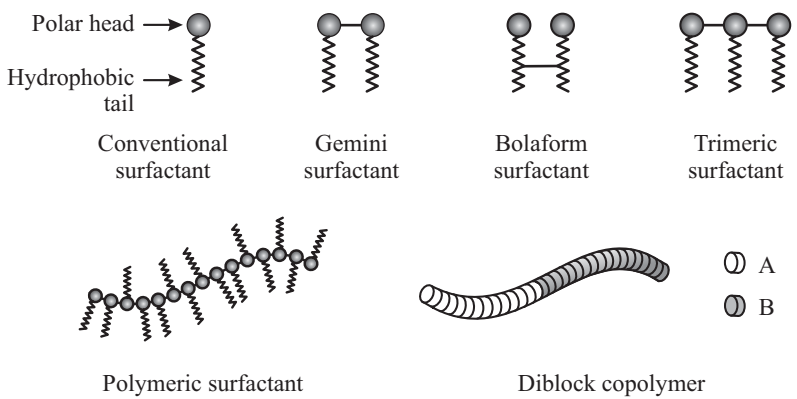



**Figure 12.1:** Different types of surface active molecules.

## 12.2 Spherical micelles, cylinders, and bilayers

### 12.2.1 The critical micelle concentration

One characteristic property of surfactants is that they spontaneously aggregate in water and form well-defined structures such as spherical micelles, cylinders, bilayers, etc. (review: Ref. [524]). These structures are sometimes called association colloids. The simplest and best understood of these is the micelle. To illustrate this we take one example, sodium dodecylsulfate (SDS), and see what happens when more and more SDS is added to water. At low concentration the anionic dodecylsulfate molecules are dissolved as individual ions. Due to their hydrocarbon chains they tend to adsorb at the air–water interface, with their hydrocarbon chains oriented towards the vapor phase. The surface tension decreases strongly with increasing concentration (Fig. 3.7). At a certain concentration, the **critical micelle concentration** or

critical micellization concentration (CMC), this decrease stops. Above the CMC, which for SDS in water is at 8.3 mM, the surface tension remains almost constant.

Similar dependencies on concentration are observed for the osmotic pressure or the electrical conductance of the solution. If we look at the optical turbidity of the solution the trend is opposite. At low concentration the solution is transparent. When the concentration reaches the CMC many solutions become opaque. In parallel, a property, which is of great practical relevance, changes: the capacity to solubilize another hydrophobic substance. At concentrations below the CMC of the surfactant, hydrophobic substances are poorly dissolved. At the CMC they start being soluble in aqueous solution. This capability increases with increasing surfactant concentration. There may be small systematic differences in the concentration at which the specific property abruptly changes and the CMC determined by different methods may be different. However, the general trend and the dependency on external parameters such as temperature or salt concentration is always the same.

What is the reason for this abrupt change at the CMC? Above the CMC the surfactants spontaneously form aggregates — the micelles . The hydrocarbon chains gather inside the aggregate and the polar head groups orient towards the aqueous phase. The result is a spherical object of typically 30–100 surfactant molecules with an oily phase inside (Fig. 12.2). Typical outer diameters are 3–6 nm as determined by light scattering, small-angle X-ray scattering (SAXS), and small-angle neutron scattering (SANS) in deuterium oxide ($D_2O$). Micelle interiors show liquid phase properties as demonstrated by nuclear magnetic resonance (NMR) [523].

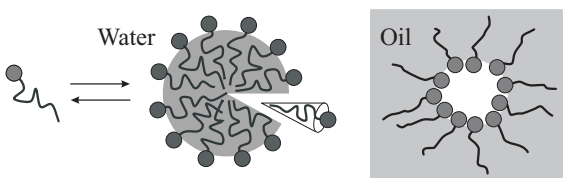



**Figure 12.2:** Schematic cross-section through a spherical micelle in water. A shell of polar head groups surrounds the hydrophobic core formed by the hydrocarbon chains. The micellar structure is in equilibrium with monomers in solution. Right: Inverted micelle in oil.

As an average a micelle contains a certain number of surfactant molecules, the mean micellar aggregation number $N_{agg}$. Not all micelles are constituted by precisely the same number of surfactants. Some micelles contain more, other micelles contain fewer surfactants so that there is a significant polydispersity. Such a distribution of aggregation numbers is schematically shown in Fig. 12.2. Slightly below the CMC, most surfactants form monomers or low number aggregates. Still, some micelles already exist. With increasing surfactant concentration the concentration of monomers increases. Also the concentration of micelles grows but only slightly. This changes at the CMC. Once the CMC is reached the monomer concentration remains almost constant and an added surfactant forms new micelles. The distribution of aggregation numbers is near Gaussian with a standard deviation $\Delta N_{agg}$, which is approximately given by the square-root of the mean aggregation number: $\Delta N_{agg} \approx \sqrt{N_{agg}}$ [109].

The situation is more complex. With increasing concentration the mean aggregation number is not strictly constant. It increases slightly with increasing total concentration.

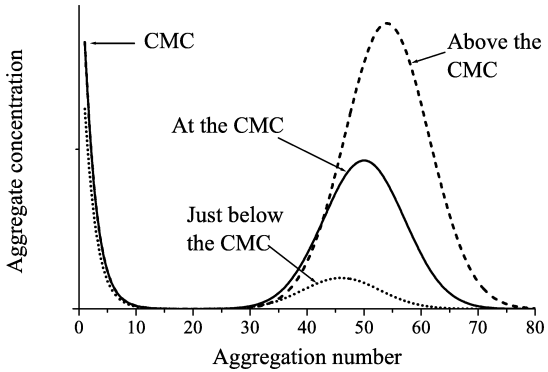



**Figure 12.3:** Schematic distribution of surfactant aggregates as a function of the aggregation number for three different concentrations. When the total surfactant concentration is equal to the CMC we assumed a mean aggregation number of $N_{agg} = 50$.

A micelle is a dynamic structure. Surfactants leave the micelle and go into solution while other surfactants enter the micelle from solution. The timescales involved depend critically on the specific structure of the surfactant, in particular on the length of the hydrocarbon chain. For example, the residence time of a single dodecylsulfate ($CH_3(CH_2)_{11}OSO_3^-$) in a SDS micelle at 25°C is 6 $\mu$s [525]. If we reduce the chain length by two methylene units to decyl sulfate ($CH_3(CH_2)_9OSO_3^-$) the residence time decreases to roughly 0.5 $\mu$s. Tetradecyl sulfate ($CH_3(CH_2)_{13}OSO_3^-$), which has two methylene units more than dodecylsulfate, typically remains 83 $\mu$s in a micelle.

### 12.2.2 Influence of temperature

For ionic surfactants micellization is surprisingly little affected by temperature considering that it is an aggregation process; later we see that salt has a much stronger influence. Only if the solution is cooled below a certain temperature does the surfactant precipitate as hydrated crystals or a liquid crystalline phase (Fig. 12.4). This leads us to the **Krafft temperature**[1] also called Krafft point [526]. The Krafft temperature is the point at which surfactant solubility equals the critical micelle concentration. Below the Krafft temperature the solubility is quite low and the solution appears to contain no micelles. Surfactants are usually significantly less effective in most applications below the Krafft temperature. Above the Krafft temperature, micelle formation becomes possible and the solubility increases rapidly.

Nonionic surfactants tend to show the opposite temperature effect: As the temperature is raised, a point may be reached at which large aggregates precipitate out into a distinct phase. The temperature at which this happens is referred to as the **cloud point**. It is usually less sharp than the Krafft temperature. [2] The phenomenon that nonionic surfactants become less soluble at elevated temperature will be important when we discuss the phase behavior of emulsions.

[1] Friedrich Krafft, 1852–1923. German chemist, professor of organic and physical chemistry in Heidelberg.

[2] This behavior is similar to that observed for polyethylene oxide in water. With increasing temperature water becomes a less good solvent for polyethylene oxide.

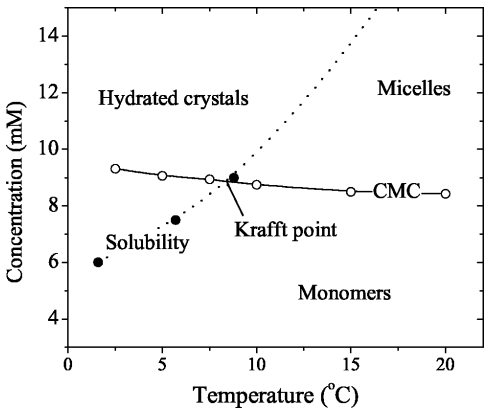



**Figure 12.4:** Solubility and CMC versus temperature for SDS in water according to Ref. [527].

### 12.2.3 Thermodynamics of micellization

There are several approaches to derive the Gibbs free energy of micellization. We only discuss one of them which is called the phase separation model. Even this approach only leads to approximate expressions for nonionic surfactants. More detailed discussions of the thermodynamics of micellization can be found in Refs. [3, 528, 529].

In the phase separation model we take advantage of the fact that micellization has much in common with the formation of a separate liquid phase. At low concentration the chemical potential of the dissolved surfactants can be described by

$$\mu_{sur}(solvent) = \mu^0_{sur} + RT \ln [S] \tag{12.1}$$

where $\mu^0_{sur}$ is the effective standard chemical potential at dilute solution and $[S]$ is the concentration of surfactants. At $[S] = \mathrm{CMC}$ the chemical potential of a surfactant in a micelle $\mu_{sur}(micelle)$ is equal to the chemical potential of a dissolved surfactant. This directly leads to

$$\mu_{sur}(micelle) = \mu^0_{sur} + RT \ln \mathrm{CMC} \tag{12.2}$$

The molar Gibbs energy of micelle formation is the Gibbs energy difference between a mole of monomers in micelles and the standard chemical potential in dilute solution:

$$\Delta G^{mic}_m = \mu_{sur}(micelle) - \mu^0_{sur} = RT \ln \mathrm{CMC} \tag{12.3}$$

For nonionic surfactants we can use this equation to calculate the change in Gibbs free energy of micellization. For ionic surfactants the change of dissociation of charges from the head groups effects the result.

CMCs are usually below 1 M. For this reason Gibbs free energies for micellization are negative i.e. it is a spontaneous process. For example, the CMC of $C_{10}E_8$ is 1 mM. The Gibbs free energy of micellization is $\Delta G^{mic}_m = RT \ln 0.001 = -17.1$ kJ per mol surfactant at 25°C.

Micelles form because of two competing factors [530]:

- Transfer of hydrocarbon chains out of water into the oil-like interior drives micellization. This is mainly an entropic effect called the hydrophobic effect [531]. At this point the reader might hesitate: surfactants aggregated into micelles are certainly more ordered than surfactant molecules dissolved in a liquid. Therefore, their entropy in a micelle should be lower. This effect, however, is small compared to the gain in entropy of the surrounding water molecules. Around a hydrocarbon chain the water molecules are highly ordered. When all the hydrocarbon chains are hidden inside micelles, the entropy of the water increases drastically. This more than compensates for the negative aggregation entropy of the surfactants themselves. As a result of the hydrophobic effect the CMC decreases with increasing length of the alkyl chain.

- Repulsion between the polar head groups as they come closer, opposes aggregation. The origin of this lateral repulsion is complex. Head groups have to be dehydrated when they approach each other closely, which leads to a hydration repulsion. In addition, the steric effect contributes. As the head groups come closer together their thermal fluctuations become smaller because they are confined by neighboring head groups. This reduced mobility decreases their entropy and leads to a repulsion.

  For charged head groups there is an additional electrostatic repulsion because it requires energy to bring two equal charges close together. For this reason the CMCs of ionic surfactants are usually higher than CMCs of nonionic surfactants. If we add salt to the solution this has a drastic effect on micellization of charged surfactants. The CMCs of ionic surfactants decrease with increasing concentration of background salt; salt effectively screens the electrostatic repulsion between head groups. It is energetically easier to bring ionic surfactants together when the charges are effectively screened.

As an example, the following diagram shows how the CMCs of different alkylsulfates decrease with increasing background salt concentration. As a result of the reduced electrostatic repulsion due to salt the Gibbs energy of micellization also increases. For example, for SDS it goes from $\Delta G_m^{mic}/k_BT = -13.0$ without salt to $\Delta G_m^{mic}/k_BT = -14.2$ per mol surfactant at 0.3 M NaCl [528]. In addition, the CMC drastically decreases with increasing length of the alkyl chain because of the hydrophobic effect. The longer the alkyl chain the stronger is the tendency to aggregate.

Micelles are not completely hydrophilic on their surface. The lateral repulsion between the head groups is usually so strong that, in between, some hydrophobic inner part of the micelle is directly exposed to water.

■ **Example 12.1.** The CMC of $C_{12}E_7$ is 0.083 mM at room temperature. By SANS and dynamic light scattering the mean hydrocarbon core radius was found to be 1.70 nm at a surfactant concentration of 2 mM [532]. The mean aggregation number is 64. If we divide the total surface area of the core by the number of surfactants, we get the area per molecule at the core radius. It is $4\pi\,(1.7\ \mathrm{nm})^2/64 = 0.57\ \mathrm{nm}^2$. The cross-sectional area of polyethylene oxide is below 0.2 nm$^2$. So, more than half the core area is exposed to aqueous or at least to a polar medium.

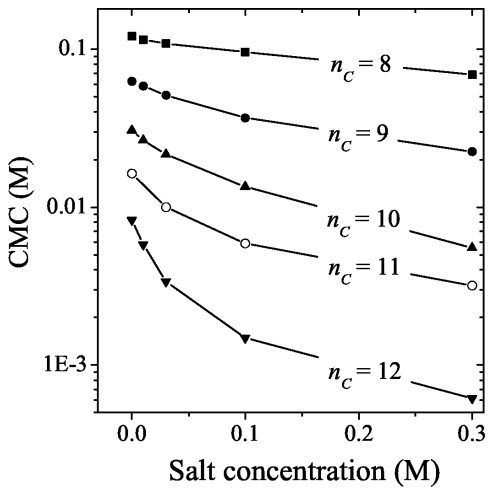



**Figure 12.5:** CMCs of different alkylsulfates (from octylsulfate, $n_C = 8$, to dodecylsulfate, $n_C = 12$) versus the concentration of additional NaCl in aqueous solution at 21°C (Ref. [4] p. 189).

### 12.2.4 Structure of surfactant aggregates

Surfactants not only aggregate to spherical micelles but also form cylinders, bilayers, inverted micelles, etc. [524]. The type of aggregate structure formed depends on different factors. An important factor is the so-called **surfactant parameter**, also referred to as the packing ratio [533]:

$$N_S = \frac{V_C}{L_C \sigma_A} \tag{12.4}$$

Here, $V_C$ is the volume of the hydrophobic part of the surfactant, $L_C$ is the length of the hydrocarbon chains, and $\sigma_A$ is the effective area per head group.

From the density of pure hydrocarbons a simple equation for the volume of a saturated hydrocarbon chain (no double bonds) can be obtained (see exercises):

$$V_C \approx (n_C \cdot 0.027 + 0.029)\,\mathrm{nm}^3 \tag{12.5}$$

The length can be estimated from

$$L_C \approx (n_C \cdot 0.127 + 0.15)\,\mathrm{nm} \tag{12.6}$$

The first number, 0.127 nm, is the carbon–carbon bond length (0.154 nm) projected onto the direction of the alkyl chain in all-*trans* configurations. The second number, 0.15 nm, is the van der Waals radius of the terminal methyl group, minus 0.127 nm.

The most problematic quantity in the definition of the surfactant parameter is the head group area. For ionic surfactants $\sigma_A$ depends on both the electrolyte and the surfactant concentration. In this case the surfactant parameter is only of limited usefulness for a quantitative

description. For nonionic or zwitterionic surfactants, $\sigma_A$ is relatively insensitive on external conditions and it is a parameter characterizing the surfactant. Still, it is very difficult to make an a priori calculation.

■ **Example 12.2.** The head group area of SDS in water (no additional background salt) is 0.62 nm$^2$. This leads to a surfactant parameter of

$$N_S = \frac{V_C}{L_C \sigma_A} = \frac{(110.027 + 0.029)\,\mathrm{nm}^3}{(110.127 + 0.15)\,\mathrm{nm} \cdot 0.62\ \mathrm{nm}^2} = 0.37 \tag{12.7}$$

The surfactant parameter relates the geometry of the molecule to the preferred curvature of the aggregates formed. Small values of $N_S$ imply highly curved aggregates. To illustrate this we first have a look at the surfactant parameters of simple geometric objects (Fig. 12.6). For a cone of height $L_C$ and a surface area $\sigma_A$ the surfactant parameter is 1/3. For a wedge of height $L_C$ and a surface area $\sigma_A$ it is 0.5, for a cylinder we have $N_S = 1$.

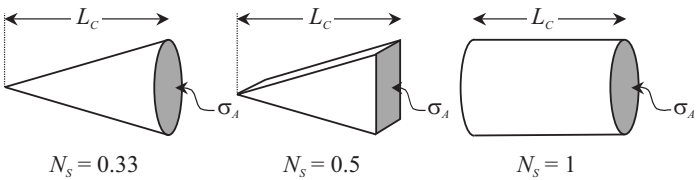



**Figure 12.6:** Surfactant numbers for a cone, a wedge, and a cylinder.

Spherical micelles are formed for $N_S \approx 0.33$. SDS, for example, tends to form spherical micelles in water. For SDS the aggregation number is 56 and the radius of the hydrophobic interior, which is roughly equal to $L_C$, is 1.7 nm. The effective head group area is 0.62 nm$^2$. This is significantly larger then the actual cross-sectional area of a sulfate group of 0.27 nm$^2$. Again we see that more than 50% of the micelles surface is hydrophobic. This clearly demonstrates that $\sigma_A$ is an effective head group area, which is determined by two opposing effects: hydrophobic attraction and lateral repulsion between head groups. Most charged surfactants form micelles at low salt concentrations because the electrostatic repulsion leads to large head group areas. With increasing salt concentration the surfactant parameter increases and other structures are formed.

Cylindrical (rod-like) micelles are formed when $N_S \approx 0.5$ (Fig. 12.7). The ends of the cylinders are capped by hemispheres to avoid exposure of the hydrocarbon interior to water. Although the diameter of the cylinders is defined by the length of the surfactant, cylindrical micelles are usually polydisperse because the cylinders can grow to varying lengths by just incorporating more and more surfactants. At high surfactant concentrations the cylinders often form into a hexagonal packing. Falling into this category are single chained surfactants with charged head groups at high salt concentration (e.g. SDS, CTAB) and those possessing nonionic (e.g. $C_{12}E_5$) or zwitterionic head groups.

■ **Example 12.3.** Cylindrical micelles formed from CTAB in aqueous medium and adsorbed to gold(111) are shown on the cover of this book. The AFM image were taken in a special non-contact mode to avoid destruction of the micelles. The cylinders orient parallel to steps in the gold.

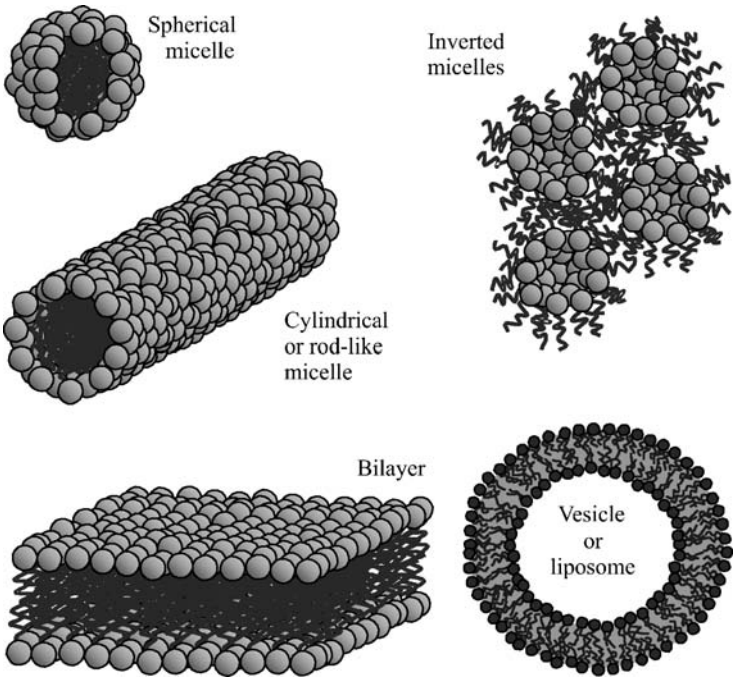



**Figure 12.7:** Aggregates formed by surfactants.

Bilayers are preferentially formed for $N_S = 0.5...1$. Lipids that form bilayers cannot pack into micellar or cylindrical structures because of their small head group area and because their alkyl chains are too bulky to fit into a micelle. For bilayer-forming lipids this requires that for the same head group area $\sigma_A$, and chain length $L_C$, the alkyl chains must have twice the volume. For this reason lipids with two alkyl chains are likely to form bilayers. Examples are double-chained phospholipids such as phophatidyl choline or phophatidyl ethanolamine. Lipids with surfactant parameters slightly below 1 tend to form flexible bilayers or vesicles. Lipids with $N_S = 1$ form real planar bilayers. At high lipid concentration this leads to a so-called lamellar phase . A lamellar phase consist of stacks of roughly parallel planar bilayers. In some cases more complex, bicontinuous structures are also formed. As indicated by the name, bicontinuous structures consist of two continuous phases.

Surfactants with very small head group areas such as cholesterol tend to form inverted micelles (Fig. 12.7). Their head groups point into the center of the micelle while the hydrophobic tails form the continuous, hydrophobic outer region. Inverted structures such as inverted liposomes, are also formed in nonpolar solvents such as toluene, benzene, cyclohexane instead of water [534].

### 12.2.5 Biological membranes

Lipid bilayers are of fundamental importance in biology. All biological membranes are formed by lipid bilayers. They separate the interior of cells from the outside world and they separate different compartments in eucaryontic cells. Why are they such ideal structures for membranes? Their main task is to avoid diffusion of polar molecules (such as sugars, nucleotides) and ions (in particular $Ca^{2+}$, $Na^{+}$, $K^{+}$, and $Cl^{-}$) from one compartment into another. The hydrophobic interior of the lipid bilayers efficiently achieves this. Polar molecules and especially ions cannot pass the hydrophobic interior. To transfer, for instance, an ion of radius $R = 2$ Å from the water phase ($\varepsilon_1 = 78$) into a hydrocarbon environment ($\varepsilon_2 = 4$) the change in Gibbs free energy is [535]

$$\frac{e^2}{8\pi\varepsilon_0 R}\left(\frac{1}{\varepsilon_2} - \frac{1}{\varepsilon_1}\right) = 5.77 \times 10^{-19}\,\mathrm{J}\left(\frac{1}{4} - \frac{1}{78}\right) = 1.37 \times 10^{-19}\,\mathrm{J} \tag{12.8}$$

This is 33 $k_B T$ and ions have practically no chance to pass the interior of a lipid membrane. As a result the electric resistivity of a lipid bilayer in aqueous electrolyte is extremely high. In general, the electrical resistivity of a membrane $R_e$ is inversely proportional to its area $A$:

$$R_e = \frac{R_{mem}}{A} \tag{12.9}$$

Here, $R_{mem}$ is the specific membrane resistance in units of $\omega\mathrm{cm}^2$. For lipid bilayers $R_{mem}$ is of the order of $10^8\,\Omega\mathrm{cm}^2$. If we built a membrane of similar thickness ($\approx 4$ nm) of a good insulator like porcelaine (specific resistivity $\approx 10^{14}$ $\Omega$cm) its membrane resistance would only be $10^{14}\,\Omega\mathrm{cm} \cdot 10^{-7}\,\mathrm{cm} = 10^7\,\Omega\mathrm{cm}^2$. In addition, a bilayer can stand potentials of typically 200 mV, which results in an enormous electric field strength of $\approx 10^8$ V/m.

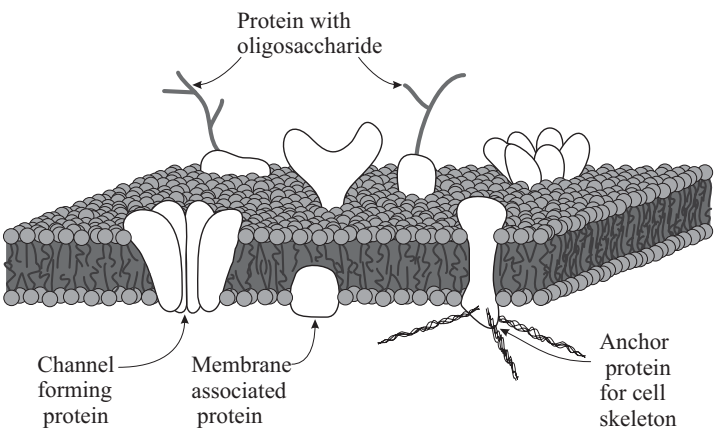



**Figure 12.8:** Schematic picture of a biomembrane with associated and transmembrane proteins. As with most schematics, some things are over simplified. For example, in reality the lipid head groups (typical diameter 0.6 nm) are very much smaller than the proteins (typical diameter 3–5 nm). In addition, lipid bilayers are flexible and are usually not perfectly planar.

If charged or polar molecules cannot pass the hydrophobic interior of a lipid biliayer, how can nutrients or ions be transported into the cell? How are signals transduced from the extracellular to the intracellular side? All these functions are performed by membrane proteins (Fig. 12.8). These membrane proteins are associated with or integrated into the lipid bilayer. There are, for instance, channel proteins which allow specific ions or water-soluble molecules to pass once a certain signal tells the channel to open. Such a signal can be the binding of a certain chemical or an electrical potential across the membrane. Other proteins actively pump molecules or ions through the membrane. Receptor proteins bind specific messenger molecules like hormones, change their conformation, and initiate a reaction on the other side of the membrane. Some membrane proteins in cell membranes have oligosaccharides attached on the extracellular side. These oligosaccharides are important for cell recognition. On the inner side specific proteins serve as anchors for the cell skeleton.

Lipid bilayers are flexible and they cannot resist significant mechanical stress. Cells which have to survive external forces such as bacteria or cells which have to provide mechanical stability have an outer cell wall in addition to the lipid membrane.

## 12.3 Macroemulsions

### 12.3.1 General properties

Oil and water do not mix. On many occasions, however, we want them to mix. This is accomplished by emulsification. An **emulsion** is a dispersion of two immiscible liquids. One phase is the dispersing agent and is called the external, outer, or continuous phase. At least one further phase is finely distributed (disperse phase). It is also called the inner or internal phase. Emulsions are of fundamental importance in many applications and various fields of science and technology such as oil recovery and the production of creams for the pharmaceutical and cosmetic industry. In cooking and in food industry emulsions have numerous uses in products such as margarines, soups, sauces, chocolate drinks, etc. [536, 537].

From the thermodynamic point of view, we can distinguish two types of emulsion. Thermodynamically stable emulsions are called microemulsions. Metastable (or unstable) systems are known as macroemulsions. Microemulsions typically involve smaller length scales than macroemulsions, which is indicated by the prefixes. While droplets in a microemulsion are 5–100 nm in diameter, a typical diameter in a macroemulsion is of the order of the wavelength of visible light (0.5–10 $\mu$m). Macroemulsions are usually opaque because the droplets are large enough to scatter light. This is for instance the reason why milk is white. Milk consists of water in which oil droplets are dispersed. In contrast, microemulsions are transparent; the droplets are too small to scatter the light effectively.

Macroemulsions tend to separate into two phases. This is called demulsification. Demulsification can be very slow so that even a macroemulsion might appear stable. Often macroemulsions are just called emulsions because all classical dispersions of oil and water were macroemulsions. In this section we use the term "emulsion" if it concerns a property of macro- *and* microemulsions. If we discuss the properties primarily of macroemulsions we use the full term. From the practical point of view micro- and macroemulsions are very different. That we discuss some properties together should not hide this fact.

The most important emulsions are water-in-oil emulsions (W/O emulsions) and oil-in-water emulsions (O/W emulsions). Oil designates here any liquid not soluble in water. In an oil-in-water emulsion, water forms the continuous phase with dispersed drops of oil (Fig. 12.9). Milk is one example. In case oil is in the continuous phase, we have a water-in-oil emulsion. Introductions into emulsions are Refs. [4, 538, 540].

How can we experimentally determine which is the outer and which the inner phase? One possibility is to use electron microscopy which provides detailed images of the emulsion structure. Electron microscopes are relatively expensive and sample preparation requires time and skill. Therefore alternative techniques are often used:

- One way is to measure the electrical conductivity. An oil-in-water emulsion has a higher conductivity because water as the continuous phase has a much higher conductivity than oil.
- If we carefully add one of the phases to the emulsion it should only be able to mix if it is in the continuous phase. Otherwise, we expect to get a layer of the added phase plus a layer of emulsion.
- When an optically visible dye is carefully added, the color of the emulsion only changes if it is dissolved in the continuous phase. If it is not soluble in the continuous phase, it has no chance to spread.

An important quantity, which characterizes a macroemulsion, is the volume fraction of the disperse phase $\phi_d$ (inner phase volume fraction). Intuitively one would assume that the volume fraction should be significantly below 50%. In reality much higher volume fractions are reached. If the inner phase consists of spherical drops all of the same size, then the maximal volume fraction is that of closed packed spheres ($\phi_d = 0.74$). It is possible to prepare macroemulsions with even higher volume fractions; volume fractions of more than 99% have been achieved. Such emulsions are also called "high internal phase emulsions" (HIPE). Two effects can occur. First, the droplet size distribution is usually inhomogeneous, so that small drops fill the free volume between large drops (see Fig. 12.9). Second, the drops can deform, so that in the end only a thin film of the continuous phase remains between neighboring droplets.

The inner phase volume fraction determines many properties of an emulsion. One example is the viscosity $\eta_{em}$. For small volume fractions one can often regard the disperse phase as consisting of rigid, spherical particles instead of liquid, flexible drops. Then we can apply Einstein's[3] equation [541], with $\eta$ being the viscosity of the pure dispersing agent:

$$\eta_{em} = \eta \cdot (1 + 2.5\phi_d) \tag{12.10}$$

The electrical conductivity of dilute emulsions can be treated by classical electrodynamic theory and the conductivity is given by

$$\kappa_e = \kappa_e^c \cdot \frac{1 + 2a\phi_d}{1 - a\phi_d} \qquad \text{with} \qquad a = \frac{\kappa_e^d - \kappa_e^c}{\kappa_e^d + 2\kappa_e^c} \tag{12.11}$$

[3] Albert Einstein, 1879–1955. German physicist, professor in Zürich, Prague, and Berlin. Emigrated to the US (Princeton) in 1933. Nobel price for physics in 1921.

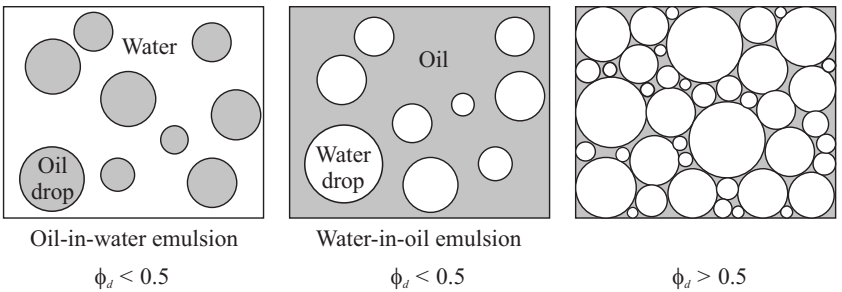



**Figure 12.9:** An oil-in-water emulsion and a water-in-oil emulsion with low volume fractions of the dispersed phase $\phi_d$. In addition, a water-in-oil emulsion with high volume fraction is shown. Volume fractions above 0.74 can occur due to the polydispersity of the drops. Small drops can fill the spaces between large drops.

where $\kappa_e^c$ and $\kappa_e^d$ are the conductivities of the pure continuous phase and the pure disperse phase, respectively. Equation (12.11) is only valid for low volume fractions ($\phi_d \leq 0.2$), where the spacing between spherical drops is large. For higher volume fractions more complicated expressions have to be used [539].

Practically, no emulsion is monodisperse and drops of the inner phase have different sizes. To characterize their size distribution, the following log-normal distribution function has proved to be useful:

$$P = \frac{1}{\Delta R\sqrt{2\pi}} \exp\left[\frac{(\ln R - \ln \bar{R})^2}{2 \cdot \Delta R^2}\right] \tag{12.12}$$

$P(R)dR$ is the probability of finding a droplet with a radius in the interval $R \ldots R + dR$, $\bar{R}$ is the mean radius and $\Delta R$ is the standard deviation. $\bar{R}$ can be determined by light scattering.

### 12.3.2 Formation

If two pure, immiscible liquids, such as decane and water, are vigorously shaken together, they will form a dispersion. When stopping the agitation, phase separation occurs so quickly that it is questionable whether the term emulsion really should be applied to the system. The driving force for phase separation is the surface tension between the two phases. Upon coalescence the total interfacial area is reduced, which leads to a lower Gibbs free energy of the system. A surfactant component is generally needed to obtain a reasonably stable macroemulsion. If, for instance, a little soap is added to the water–decane system, the result of shaking is a true macroemulsion that separates out very slowly. The minimal work required to disperse a liquid of volume $V$ in a solvent is

$$\Delta G_{em} = \gamma \frac{3V}{R} \tag{12.13}$$

Here, $R$ is the radius of the droplets and $\gamma$ is the liquid–liquid interfacial tension (see Section 13.5 and Table 13.2). The amount of $\Delta G_{em}$ is equal to the Gibbs free energy gained upon demulsification.

■ **Example 12.4.** Dispersion of 0.1 L of water in oil into droplets of 100 nm diameter requires at least an energy of 246 J, assuming an interfacial tension for oil-water of 41 mN/m. In total, $\approx 2 \times 10^{17}$ droplets are formed. The molar Gibbs energy of emulsification is $\Delta G_{em} = 44$ J/mol using a molar volume of 18 cm$^3$. This is much lower than typical chemical bonding energies, which are of the order of several $10^4$ J/mol. Increasing the droplets or adding surfactant (to decrease $\gamma$) reduces $\Delta G_{em}$.

In practice, we have to put in much more energy. There are some examples of how chemical energy can be used to form macroemulsions (Ref. [4], p. 573). Usually, however, macroemulsions are formed by mechanical agitation and most energy is converted to heat through viscous dissipation. During mechanical agitation not only large drops are separated to form smaller droplets but also already existing small droplets collide and coalesce again to larger drops. In addition, an energy barrier has to be overcome during the formation of smaller droplets from bigger ones.

To form macroemulsions, different mechanical devices are used such as homogenizers, stirrers, or ultrasonic transducers. Alternatively, one liquid can be pressed through a filter with defined pore size into the immiscible liquid. One important question is how the average droplet size varies with the energy input for different emulsification techniques. We begin our qualitative discussion by realizing that any splitting of drops into smaller droplets is preceded by a deformation. A deformation of a drop of radius $R$ is resisted by the internal pressure of the drop, which is the Laplace pressure $\Delta P = 2\gamma/R$. To tear such a drop into two parts, we need a pressure gradient. The required externally applied pressure gradient is proportional to the Laplace pressure divided by a typical size of the droplet $R$:

$$\frac{\Delta P_{ext}}{\Delta x} \propto \frac{2\gamma}{R^2} \tag{12.14}$$

The smaller the drop size the larger the pressure gradient needs to be. We can also see, that surfactants facilitate emulsification because they reduce the surface tension.

If we analyze different mechanical emulsification techniques and relate the pressure gradient to the work required $W$ (energy input) we find that for most devices the mean radius of drops $\bar{R}$ scales with

$$\bar{R} \propto W^{-0.4...0.6} \tag{12.15}$$

Macroemulsions are in a nonequilibrium state and their properties depend not only on parameters such as temperature and composition but also on the method of preparation. This leads to a high level of complexity when it comes to scientific experimental studies or practical applications. Recipes which work at one place often do not lead to the same result in another laboratory, because seemingly insignificant details have a big effect. How oil and water are mixed if, for instance, they are shaken or swirled or if air is bubbled through, and the wetting behavior of the vessel, can change the outcome dramatically.

### 12.3.3 Stabilization

To stabilize a macroemulsion, the dispersed phase should be prevented from reaching molecular contact. Emulsions are stabilized by different agents, which all adsorb strongly at the

liquid–liquid interface. They are called emulsifiers. It is not only surfactant that can be effective emulsifiers. For industrial applications macromolecules, like proteins, polysaccharides or synthetic polymers are important. One reason for their stabilizing effect is that, due to their size, they adsorb practically irreversibly at the interface. Polymers with weakly surface active monomers adsorb at the oil–water interface and the polymer chains will go into both media. An effective polymeric emulsifier should prefer the continuous medium, to generate a strong steric repulsion (Fig. 12.10).

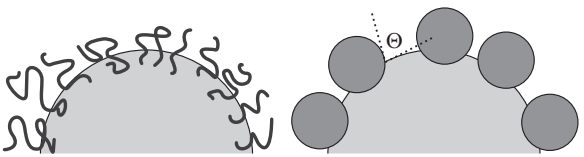



**Figure 12.10:** Droplet stabilized by polymer (left) and by adsorbed solid particles (right). The contact angles $\Theta$ of the solid particles with the continuous phase should be smaller than $90°$.

Powders often have a stabilizing effect on emulsions [548]. To understand the responsible effect we have to remember that a particle assumes a stable position in the liquid–liquid interface if the contact angle is not zero (see section 7.2.2). Upon coalescence of two drops the solid particles would have to desorb from the interface. This is energetically unfavorable. A common example of the stabilizing contribution of solid particles are margarine and butter. Both are water-in-oil emulsions. The water droplets are stabilized by small fat crystals.

Whether, upon stabilizing with powders, an oil-in-water or a water-in-oil emulsion is formed depends largely on the contact angle. Also the effect of additional surfactants can often be explained by their influence on the contact angle.

■ **Example 12.5.** The stabilizing effect of powders was impressively demonstrated by making liquid marbles in air [549]. Liquid marbles (Fig. 12.11) are obtained by making a small amount of water (typically 1 $mm^3$) roll on a very hydrophobic powder. The powder particles go into the interface and completely coat it so that, after spontaneous formation of the spherical drop, only the solid caps of powder particles come into contact with the solid support.

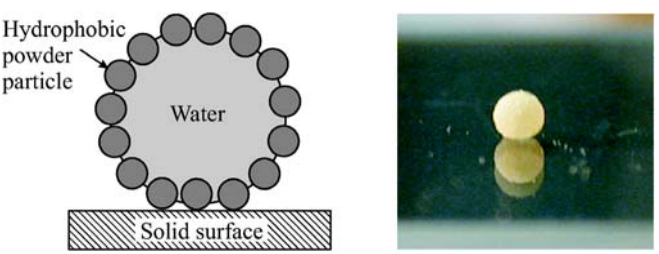



**Figure 12.11:** Liquid marble on a planar solid surface stabilized by a hydrophobic powder. A schematic (left) and a light microscope image (right) are shown. On the right the marble and its mirror image can bee seen. Thanks to David Quéré for providing us with the figure.

Some inorganic electrolytes stabilize oil-in-water emulsions. One example is potassium thiocyanide (KCNS), which dissociates in the aqueous phase. The anion $CNS^-$ adsorbs at the interface, which becomes negatively charged. As a result the oil droplets repel each other electrostatically.

The most common agents to stabilize an emulsion are surfactants. Different effects contribute to the stabilization of emulsions. Steric repulsion between those parts of the surfactant, which are in the continuous phase, is an important effect. For a water-in-oil emulsion the hydrocarbon chains are hindered in their thermal movements if two water drops approach each other too closely. For an oil-in-water emulsion there is an additional effect: the hydrophilic head groups have to be dehydrated to come into close contact. The resulting hydration repulsion stabilizes the emulsion.

Electrostatic forces, acting when the electric double layers of two drops overlap, play an important role. As mentioned above, oil drops are often negatively charged because anions dissolve in oil somewhat better than cations. Thus, the addition of salt increases the negative charge of the oil drops (thus their electrostatic repulsion). At the same time it reduces the Debye length and weakens the electrostatic force. For this reason, emulsion stability can exhibit a maximum depending on the salt concentration.

For ionic surfactants another effect often dominates and usually salt tends to stabilize emulsions. Reason: without salt the distance between surfactants in the interface is large because the molecules electrostatically repel each other. This prevents a high surface excess. The addition of salt reduces this lateral repulsion and more surfactant molecules can adsorb at the interface. Then, according to the Gibbs adsorption isotherm Eq. (3.52), the surface tension is reduced and the emulsion is stabilized.

■ **Example 12.6.** Between paraffin oil and water, the interfacial tension at 25°C is $\gamma = 41$ mJ/m$^2$. The addition of 1 mM cis-9-octadecanoic acid (oleic acid) reduces the interfacial tension to 31 mJ/m$^2$. As a result the solution becomes acidic. On neutralization by 1 mM NaOH the interfacial tension falls to 7.2 mJ/m$^2$. Further addition of 1 mM NaCl decreases the interfacial tension even to 0.01 mJ/m$^2$ [550].

Whether the system formed on mixing oil, water, and surfactant will be an oil-in-water or a water-in-oil emulsion is a central problem in emulsion technology. It was realized very early that the volume fractions of oil and water are not that important and that the type of emulsion is primarily determined by the nature of the surfactant. Simply speaking: surfactants with $N_S < 1$ tend to form oil-in-water emulsions, while surfactants with $N_S > 1$ are more likely to form water-in-oil emulsions. Two more detailed guiding principles which are used for practical emulsion formulation are Bancroft's rule of thumb and the more quantitative concept of the HLB scale:

- Surfactants are enriched at the interface but they are also dissolved in the aqueous and the oil phase. Some surfactants are more soluble in water, others are better soluble in oil. In essence, Bancroft's rule states that the continuous phase of an emulsion will be the phase in which the emulsifier is preferentially soluble [542, 543].

- Griffin suggested an empirical quantitative hydrophile–lipophile balance (HLB) scale which characterizes the tendency of a surfactant to form water-in-oil or oil-in-water emulsions [544]. The HLB is a direct measure of the hydrophilic character of a surfactant: the

larger the HLB, the more hydrophilic the compound. For most surfactants the scale runs from 3–20. Surfactants with low HLB values (3–6) tend to stabilize water-in-oil emulsions, while those with high HLB values (8–18) tend to form oil-in-water emulsions. How HLB values are determined is difficult and would exceed the scope of this book. A detailed description is contained in Refs. [545] and [538] (p. 232).

Later we discover another parameter, the phase inversion temperature(PIT), which helps us to predict the structure of emulsions stabilized by nonionic surfactants. The PIT concept is based on the idea that the type of an emulsion is determined by the preferred curvature of the surfactant film. For a modern introduction into the HLB and PIT concepts see Ref. [546].

Up to this point we have concentrated on static effects. Often dynamic effects are equally important in stabilizing emulsions. An emulsifier, which can dissolve only slowly from the interface, has a stabilizing effect, because during the merging of two drops the total surface diminishes, i.e., emulsifier must desorb from the interface.

In several processes, for instance, in budding of a droplet from a larger one or directly after merging of two drops, the surfactant is not distributed homogeneously at the interface. Then the Marangoni effect (see Section 3.4.4) comes into play. Due to the nonuniform surfactant distribution, those regions with a low surfactant concentration have a high surface tension and vice versa. Surfactant will flow toward the area of low surfactant concentration and high surface tension. This Marangoni effect quickly establishes a uniform distribution of surfactant again. It is usually faster than diffusion equilibration.

### 12.3.4 Evolution and aging

Freshly prepared macroemulsions change their properties with time. The time scale can vary from seconds (then it might not even be appropriate to talk about an emulsion) to many years. To understand the evolution of emulsions we have to take different effects into account. First, any reduction of the surface tension reduces the driving force of coalescence and stabilizes emulsions. Second, repulsive interfacial film and interdroplet forces can prevent droplet coalescence and delay demulsification. Here, all those forces discussed in Section 6.5.3 are relevant. Third, dynamic effects such as the diffusion of surfactants into and out of the interface can have a drastic effect.

Evolution and eventually demulsification of macroemulsions proceeds through a series of steps. For different macroemulsions the individual steps might be different. Rather typical is that the dispersed drops first form loose clusters without losing their identity. This process is called flocculation[4] (Fig. 12.12). It is caused by the secondary energy minimum which we discussed in Section 6.5.3. In fact, for oil-in-water emulsions the same interactions are present: attractive van der Waals forces between oil droplets destabilize the emulsion, the electrostatic double-layer repulsion stabilizes it. For water-in-oil emulsions the interactions are different. Since oil has a low dielectric permittivity only few ions are dissolved (see Eq. 12.8) and there is no significant double-layer repulsion. Still, a secondary energy minimum might exist. In the secondary energy minimum the surface films are not in direct molecular contact because the repulsive energy maximum prevents such contact.

[4] In chemical engineering and mineral processing the term flocculation is used for aggregation induced by the addition of polymer while coagulation indicates aggregation induced by the addition of electrolyte.

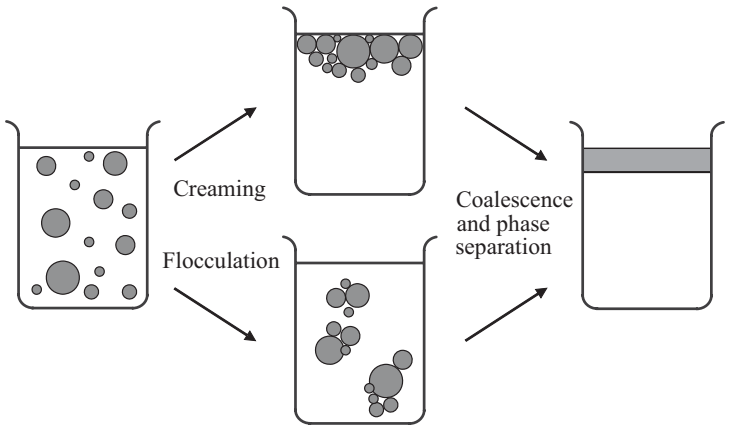



**Figure 12.12:** Typical steps in the evolution of a macroemulsion.

**Flocculation kinetics** can be described in different ways. Here we introduce a treatment first suggested by Smoluchowski [547], and described in Ref. [538], p. 417. The formalism can also be used to treat the aggregation of sols. A prerequisite for coalescence is that droplets encounter each other and collide. Smoluchowski calculated the rate of diffusional encounters between spherical droplets of radius $R$. The rate of diffusion-limited encounters is $8\pi DRc^2$, where $c$ is the concentration of droplets (number of droplets per unit volume). For the diffusion coefficient $D$ we use the Stokes–Einstein relation $D = k_BT/6\pi\eta R$. The rate of diffusion-limited encounters is, at the same time, the upper limit for the decrease in droplet concentration. Both rates are equal when each encounter leads to coalescence. Then the rate of encounters is given by

$$\frac{dc}{dt} = -8\pi DRc^2 = -\frac{8\pi Rk_BT}{6\pi\eta R}c^2 = -\frac{4k_BT}{3\eta}c^2$$

or

$$\frac{dc}{dt} = -k_fc^2 \qquad \text{with} \qquad k_f = \frac{4k_BT}{3\eta} \tag{12.16}$$

The rate constant of flocculation $k_f$ depends only on the viscosity of the solution and is independent of the size of the droplets.

Equation (12.16) is only valid for droplets of identical size. Fortunately, since it does not depend on the droplet size, the flocculation rate does not change dramatically if we consider that aggregates are formed and not all droplets have the same size. To take aggregate formation into account we just have to change the flocculation rate constant to $k_f \approx 8k_BT/3\eta$.

■ **Example 12.7.** For water at 25°C ($\eta = 8.91 \times 10^{-3}$ kgm$^{-1}$s$^{-1}$) the rate of flocculation is $k_f = 6.2 \times 10^{-19}$ m$^3$s$^{-1}$. This rate constant is valid if we give the concentration in number of droplets (or particles) per cubic meter. If we report the concentration in moles of droplets per liter (M) the rate constant is $k_f = 3.7 \times 10^8$ M$^{-1}$s$^{-1}$. Taking heteroflocculation into account, the rate is twice as high.

If an energy barrier $E^*$ for coalescence exists, the rate of effective encounters is reduced. We can take this into account by adding a Boltzmann factor:

$$\frac{dc}{dt} = -k_r \cdot e^{-\frac{E^*}{k_B T}} \cdot c^2 \tag{12.17}$$

This differential equation is solved by

$$\frac{1}{c} = \frac{1}{c_0} + k_f^* t \qquad \text{with} \qquad k_f^* = k_f \cdot e^{-\frac{E^*}{k_B T}} \tag{12.18}$$

Here, $c_0$ is the original concentration of droplets at the beginning. The quantity $d\left(1/c\right)/dt$, which should be equal to $k_f$ or $k_f^*$, is used as a measure of the initial flocculation rate.

Whether flocculated or not, drops can migrate in the gravitational field and increase the concentration either at the bottom or at the top of the vessel. The latter effect is known as **creaming**. Once drops have been brought into close proximity they can coagulate into the primary energy minimum. There the surface films get into direct molecular contact, but they might still keep their identity as separate droplets. Eventually the surface film ruptures and the two droplets merge into a bigger one. This eventually leads to coalescence and demulsification.

### 12.3.5 Coalescence and demulsification

Let us turn to the final step which is coalescence. We can identify three crucial steps in the process of droplet coalescence (Fig. 12.13):

- For two drops to coalesce, the two surfactant films must first come into molecular contact. Due to the, always present, van der Waals attraction, two neighboring droplets deform, giving a flat contact area. Depending on the repulsive forces between the two surfactant films the lamella is more or less stable.

- The two surfactant films must fuse, forming a neck with direct contact between the dispersed liquid in the two droplets. Clearly, if the two surfactant films are poorly developed the lamella will rupture easily, while a fully saturated film could resist the fusion process. Fluctuations of the surfactant surface density can trigger the fusion process. Temporarily bare spots attract each other and can break the lamella locally.

- The neck must grow in size so that the two drops eventually merge completely. In this process the surfactant film can remain intact but its area and curvature changes.

Any of the three steps can be rate-determining depending on the specific nature of the system. For a detailed discussion see Ref. [551].

In many applications, our goal is to destabilize an emulsion. **Demulsification**, with the resulting two separate phases, is in fact a common task in many industrial applications (see, for example, Table 11.1). Mechanical methods are most common for demulsification. Centrifugation, for instance, leads to creaming. Gentle stirring often accelerates coalescence. In some cases the addition of a cosurfactant or salt might destabilize an emulsion. As we see in the next section, emulsions are stable only if the spontaneous curvature of the surfactant agrees with the actual radius of the droplets. A cosurfactant with a different spontaneous curvature from the main surfactant, might lead to an unstable situation.

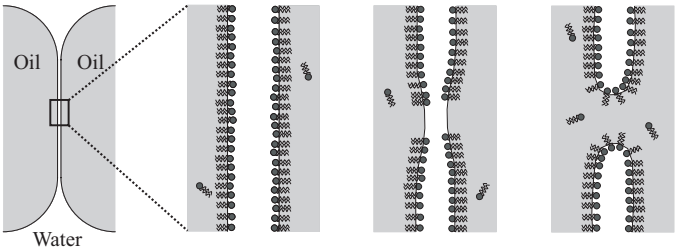



**Figure 12.13:** Three steps of droplet coalescence after flocculation or creaming for an oil-in-water emulsion.

## 12.4 Microemulsions

Thermodynamically stable emulsions were discovered some 50 years ago [552–555] and they were treated as a very special case for a long time. Thermodynamically stable not only refers to the fact that they do not change with time. It also implies that microemulsions react reversibly to changes of temperature or composition. Microemulsions tend to form spontaneously. No powerful stirring or strong agitation is required. Also the size of the structures formed does not depend on the energy input as for macroemulsions (see Eq. (12.15)). In a microemulsion, the volume fraction of surfactant is usually significant. A review about microemulsions is Ref. [556].

### 12.4.1 Size of droplets

We start our description of microemulsions by asking: what is the most likely radius of drops $R$ for given volume fractions of continuous phase, dispersed phase $\phi_d$, and surfactant $\phi_s$? This radius can be estimated by using two equations. First, the total volume of the dispersed phase is given by

$$V\phi_d = n\frac{4}{3}\pi R^3 \tag{12.19}$$

Here, $n$ is the number of droplets in the total volume $V$. For the second equation we assume that all surfactant goes into the interface and that only a negligible amount is dissolved in the continuous and dispersed phase. In addition, we assume that the surfactant is oriented normal to the interface so that the thickness of the surfactant film is equal to the total length of a surfactant molecule $L_s$. Then we get the volume occupied by surfactant:

$$V\phi_s = n \cdot 4\pi R^2 \cdot L_s \tag{12.20}$$

Dividing Eq. (12.19) by Eq. (12.20) leads to

$$R = \frac{3L_s\phi_d}{\phi_s} \tag{12.21}$$

This is the radius of droplets of the dispersed phase. It is a necessary condition for an emulsion to be thermodynamically stable. It turns out that the parameter $L_s$ should be considered as

an effective length of a surfactant molecule which can differ from the geometric length by as much as a factor of two.

■ **Example 12.8.** Consider an oil-in-water emulsion with a volume fraction of surfactant of 10% and a volume fraction of oil of 30% ($\phi_s = 0.1$, $\phi_d = 0.3$). For a surfactant with a length of 2 nm the radius of oil droplets is $R = 18$ nm. For a water-in-oil emulsion with the same parameters we have $\phi_d = 0.6$ and $R = 36$ nm.

Droplets in microemulsions are usually quite small so that the length of the surfactant is often not negligible. The complication arises as to what precisely is the radius. It can be significant whether we choose the radius to include the whole surfactant rather than, for instance, only the tail. The conventional choice is to position the radius at the so-called neutral surface, that is the plane whose area is constant upon bending. This neutral surface is usually located close to the plane where the head groups meet the hydrocarbon chains. For this reason it is chosen to be at the core radius $R_C$ (Fig. 12.14).

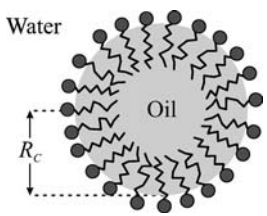



**Figure 12.14:** Schematic of an oil-in-water droplet stabilized by surfactant.

## 12.4.2 Elastic properties of surfactant films

Surfactants form semiflexible elastic films at interfaces. In general, the Gibbs free energy of a surfactant film depends on its curvature. Here we are not talking about the indirect effect of the Laplace pressure but a real mechanical effect. In fact, the interfacial tension of most microemulsions is very small so that the Laplace pressure is low. Since the curvature plays such an important role, it is useful to introduce two parameters, the principal curvatures

$$C_1 = \frac{1}{R_1} \qquad \text{and} \qquad C_2 = \frac{1}{R_2}, \tag{12.22}$$

which are the inverse of the two principal radii of curvature (see Section 2.3.1). The curvature can be positive or negative. We count it as positive if the interface is curved towards the oil phase.

Following Helfrich we can express the Gibbs free energy of curvature by an integral over the area considered [557]:

$$G = \int \left[ \frac{1}{2} k \left( C_1 + C_2 - C_0 \right)^2 + \bar{k} C_1 C_2 \right] dA \tag{12.23}$$

$C_0$ is called the spontaneous curvature. The spontaneous curvature is a more general parameter than the surfactant parameter $N_S$, defined by Eq. (12.4). It makes it easier to discuss the phase behavior of microemulsions because we get away from the simple geometric picture.

The parameters $k$ and $\bar{k}$ have the dimensions of energy and are the bending rigidity (also called bending elastic modulus) and the saddle-splay modulus (also called the modulus of Gaussian curvature), respectively. This equation has played a pivotal role in our understanding of mono- and bilayers [558]. The film is characterized by a stable equilibrium curvature, corresponding to a minimum in the Gibbs free energy, at $C_1 = C_2 = C_0 k/(2k + \bar{k})$.

■ **Example 12.9.** We consider a film with zero spontaneous curvature ($C_0 = 0$). What is the elastic energy for bending such a film to a sphere of radius $R$? With $C_1 = 1/R$ and $C_2 = 1/R$ we obtain

$$\begin{aligned} G &= \int \left[ \frac{1}{2} k \left( \frac{2}{R} \right)^2 + \frac{\bar{k}}{R^2} \right] dA \\ &= 4\pi R^2 \left[ \frac{2k}{R^2} + \frac{\bar{k}}{R^2} \right] = 4\pi \left( 2k + \bar{k} \right) \end{aligned}$$

It is independent of the radius! It does not depend on the radius because the bending energy per unit area increases but the surface area decreases. Both effects just compensate each other. See also exercise 6.

What are typical values for the parameters $k$ and $\bar{k}$? The bending rigidity of surfactant films is typically of the order of 1–20 $k_B T$ at room temperature. Factors that reduce $k$ are short alkyl chains, cosurfactants, double-chain surfactants with unequal chains, and cis-unsaturated bonds. For the saddle-splay modulus only few measurements have been done. It tends to be negative with an amount much smaller than the bending rigidity for the same system.

The spontaneous curvature can vary in the range of $-0.5$ nm$^{-1}$ to $+0.5$ nm$^{-1}$, depending on the polar head group, the length and number of apolar chains, and the nature of the oil. Most microemulsions that have been studied contain four or even more components. In addition to water, oil, and surfactant usually a cosurfactant is added. Typical cosurfactants are alcohols. With their small head group and relatively large hydrophobic tail they tend to decrease the spontaneous curvature of surfactants.

The spontaneous curvature of a surfactant film determines, to a high degree, the structure of a microemulsion. Let us illustrate this for an oil-in-water microemulsion and let us neglect the influence of the saddle-splay modulus. When the spontaneous curvature of the film is equal to the curvature of the oil droplets, $C_0 = 2/R$, we expect a stable situation [559]. When the spontaneous curvature is larger than $2/R$, the film can relax toward a lower Gibbs energy by decreasing the drop size and expelling emulsified oil to the bulk phase. When $C_0$ is much lower than $2/R$, the film relaxes by forming larger aggregates. These aggregates would have to be nonspherical because of Eq. (12.21).

### 12.4.3 Factors influencing the structure of microemulsions

What are the most important factors influencing the type of microemulsion? Here again we have to distinguish between nonionic and ionic surfactants. For nonionic surfactants, often alkylethylene glycols, temperature is the dominating parameter for the structure of a microemulsion. For ionic surfactants, mostly SDS or CTAB, the salt concentration dominates

the phase behavior. Alkyl polyglycosides take an intermediate position, although they are nonionic. Here we only discuss nonionic surfactants to show some principles.

Let us start with a generalized statement. Microemulsions stabilized with nonionic surfactants tend to form an oil-in-water type at low temperature while at high temperature water-in-oil microemulsions are more common. A phase inversion temperature exists at which the microemulsion changes from oil-in-water to water-in-oil type. This phase inversion temperature (PIT) is an important parameter characterizing an microemulsion system [560]. The reason for this phase inversion becomes clear when recalling that the tendency of the head group to bind water decreases with temperature (see end of Section 12.2.2). As a result the size of the head group area decreases. In contrast, the hydrocarbon chains tend to occupy a larger volume since their thermal fluctuations become more and more vivid. As a result, the spontaneous curvature changes from favoring an oil-in-water emulsion to favoring a water-in-oil microemulsion.

To illustrate this behavior, a phase diagram of a water-octane-$C_{12}E_5$ emulsion is shown in Fig. 12.15. We will see that the real situation is a little bit more complicated than our generalized statement at the beginning. Let us work through the phase diagram, starting from the bottom along the arrow at constant surfactant volume fraction of $\phi_s = 0.15$. For low temperature, say at 20°C, the system contains two phases: an oil-in-water microemulsion containing small oil drops coexists with an oil phase on top. The drop size is determined by the spontaneous curvature of the surfactant. Why is not all oil emulsified? Then Eq. (12.21) — a requirement for a stable phase — would be violated.

With increasing temperature the oil drops become larger and larger because the spontaneous curvature decreases. Accordingly, the volume of the oil phase decreases until at 23°C all oil is incorporated into relatively large oil drops in water and we reach a one-phase region (often denoted $L_1$) of an oil-in-water microemulsion. This point is ideally suited to determine the spontaneous radius of curvature because, here, the radius of the oil drop can be calculated from the added volume fractions with Eq. (12.21). Raising the temperature decreases $C_0$ further and from 29°C on lamellar structures are formed ($L_\alpha$). At 32°C the PIT is reached and the spontaneous curvature is zero.

Going higher in temperature results in a negative curvature (towards the hydrophilic head groups) and the surfactants start to favor water-in-oil emulsions. Now the whole sequence is repeated only with water drops in oil. Between 35°C and 37°C we have large water drops which decrease in size with increasing temperature ($L_2$). Then, above 37°C, two phases appear: a water phase underneath a water-in-oil microemulsion.

In Fig. 12.15, the phase inversion temperature is indicated by the horizontal line at $T =$ 32°C. When we move at constant temperature $T = \mathrm{PIT}$ from low to high surfactant volume fractions, we cross a three-phase region between $\phi_S \approx$1% and 5%. Here, an aqueous phase (with some surfactants dissolved as monomers), an oil phase (also with a few monomers), and a bicontinuous microemulsion coexist. With increasing $\phi_S$ the oil and water phase shrink and more and more volume is occupied by the bicontinuous microemulsion. Because of the shape of the three-phase region, the phase diagram is sometimes called a "fish cut" [562,563]. Above $\phi_S$ =5 % we enter a single-phase region consisting only of a bicontinuous microemulsion. From a practical point of view this is an important point, sometimes denoted by $\tilde{\phi}$, because it is a measure of the efficiency of a surfactant. The smaller $\tilde{\phi}$ the more efficiently a surfactant is able to stabilize an emulsion. When raising $\phi_S$ above 7% we enter the lamellar phase $L_\alpha$.

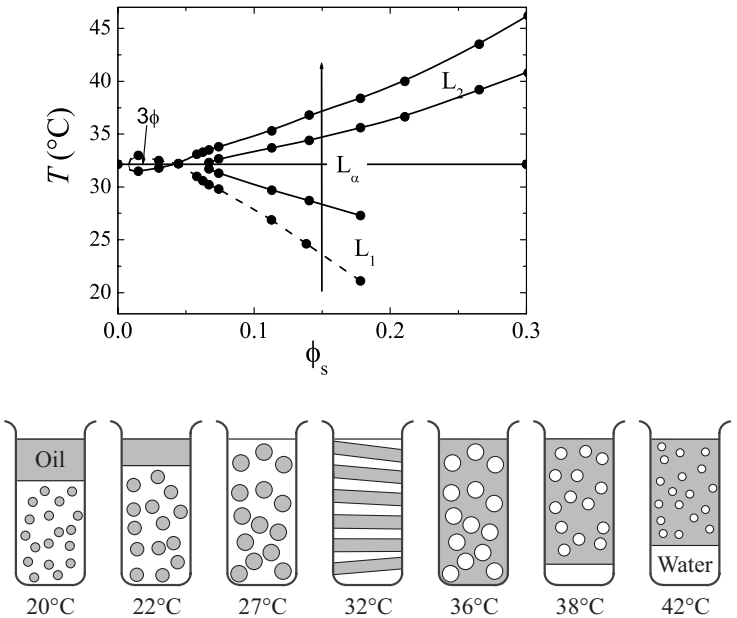



**Figure 12.15:** Phase diagram of a water-in-octane-$C_{12}E_5$ emulsion. The axes are temperature and volume fraction of surfactant. The phases are indicated by $3\phi$ (three-phase-region), $L_1$, $L_2$, and $L_\alpha$ (lamellar). The phase diagram was determined for a volume ratio between water and octane of 1:1 [561]. The phases observed along the vertical arrow at $\phi_S = 0.15$ are shown schematically at the bottom. Results were obtained by D. Vollmer.

## 12.5 Foams

### 12.5.1 Classification, application and formation

A foam is a dispersion of a gas in a liquid or a solid medium. Some properties of foams resemble those of emulsions, which is not surprising because a foam can be thought of as an emulsion in which the dispersed phase is a gas. Foams are, to our knowledge, never thermodynamically stable, they can only be kinetically trapped. This may be one reason why a quantitative description of foams is only poorly developed. Very good introductions into the subject are Refs. [564–566].

Foams have a large variety of applications. Solid foams are widely used as insulating materials. Due to the presence of air bubbles they have a low thermal conductivity. Polyurethane foams and Styrofoam are examples. Styrofoam is also used as a packing material. The light weight of polymer foams makes them attractive as filling materials to stabilize otherwise hollow structures. A natural solid foam is pumice stone. Metal foams are used in the automotive and aerospace industry as light and stable materials [567]. Ceramic foams are developed for electronic applications as piezoelectric transducers and low dielectric constant substrates [568].

To make a solid foam we start from a liquid foam and induce solidification. This can be achieved by a chemical polymerization (Styrofoam), by lowering the temperature (pumice stone or a soufflé), or by increasing the temperature to induce a structural transition (baking of bread). Porous solids can appear as solid foams because of their low density and their high content of gas. The difference between the two is that in a porous solid we have a bicontinuous structure while the individual cavities in a foam are closed. This is an important difference because porous solids tend to adsorb liquids due to capillary effects and then completely change their properties.

■ **Example 12.10.** Polyurethane (PUR) foams are widely used in the building industry. The name refers to the fact that its monomer contains urethane (∼O-CO-NH∼) as a functional group. The polymer is usually made from diisocyanates. Isocyanates are molecules which have a ∼N=C=O group attached. Upon addition of water the reaction shown in Fig. 12.16 starts.

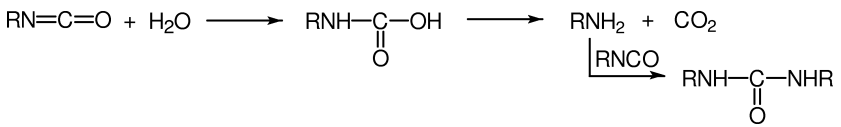



**Figure 12.16:** Reaction leading to a polyurethane foam.

Diisocyanites have two reacting groups and they can form chains or even networks. The two most commonly used diisocyanites are toluenediisocyanites (TDI) and diphenylmethane diisocyanates (MDI) (Fig. 12.17). With each bond formed one molecule of gaseous carbon dioxide is released. This is responsible for the foaming effect.

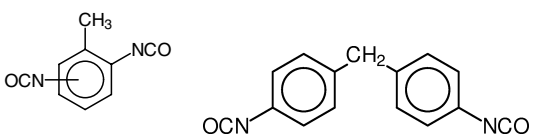



**Figure 12.17:** Structure of TDI (left) and MDI (right).

The basic step in the formation of a liquid foam is the generation of bubbles. We can accomplish this by forcing gas through a nozzle. We can also nucleate bubbles from supersaturated solutions, such as in beer, or from a superheated liquid (boiling). Another way to form bubbles is by mechanical agitation such as in a washing machine or in breaking waves on the sea.

The first application which comes into mind when thinking about liquid foams is in cleaning processes, such a shampoo or shaving foams. More important, however, are liquid foams in mineral froth flotation and, a closely related process, the de-inking of recycled paper. Other uses include fire fighting. In food we often find foams such as whipped cream or egg whites.

Foams, or more generally the formation of bubbles, can also be a nuisance. For example, the mixing of two liquids in an industrial process can be significantly slowed down by the development of foam. In the glass industry, the release of gas from the melt can lead to the formation of bubbles in the final glass.

### 12.5.2 Structure of foams

We distinguish two foam structures. For a low-volume fraction of gas the individual bubbles are spherical and are separated from each other by thick films. Such a foam is called *Kugelschaum*, according to Manegold [569], which comes from the German words sphere (*Kugel*) and foam (*Schaum*). The second type of foam contains mostly gas phase separated by thin films or lamellas. The individual cells are polyhedral in shape and the foam can be thought of as a space-filling packing of polyhedra. Such foams are sometimes called *Polyederschaum*. Usually polyhedral foams are formed by sufficient drainage from *Kugelschäumen* (Fig. 12.18).

Let us discuss the structure of a metastable polyhedral foam in a bit more detail. In pioneering experimental studies Joseph Plateau[5] established some simple rules in the second half of the $19^{th}$ century. Three of these

rules are:

- Three flat sides of polyhedra meet at an angle of 120°.
- If four or more sides come together in one line this is an unstable configuration.
- At all corners of polyhedra the four edges meet in a tetrahedral arrangement. The angle between any two edges is 109.5°.

When the number and volume of the polyhedral compartments are given, the optimal structure of the foam is the one that creates the smallest total film area. This condition constitutes a formidable but straightforward mathematical optimization problem. Solution: as an average, the polyhedra consist of 13.4 sides. Experimentally it was indeed found that the polyhedra most commonly found in foams have 14 sides, followed by 12 sides as a second choice.

### 12.5.3 Soap films

One of the central questions in the stability of foams is: why are liquid films between two adjacent bubbles stable, at least for some time? In fact, a film of a pure liquid is not stable at all and will rupture immediately. Formally this can be attributed to the van der Waals attraction between the two gas phases across the liquid. As for emulsions, surfactant has to be added to stabilize a liquid film. The surfactant adsorbs to the two surfaces and reduces the surface tension. The main effect, however, is that the surfactant has to cause a repulsive force between the two parallel gas–liquid interfaces. Different interactions can stabilize foam films [570]. For example, if we take an ionic surfactant, the electrostatic double-layer repulsion will have a stabilizing effect.

The interaction between the two gas–liquid interfaces across a foam film can directly be measured by a thin-film balance (TFB) [571–573]. A single thin foam film is formed in a hole drilled through a porous glass plate (Fig. 12.19). The plate, and therefore the liquid in the foam, is connected to a reservoir with a constant reference pressure $P_r$ by a capillary tube. The film holder is placed into a closed cell. A constant pressure $P_g$ is applied to the gas in the

[5] Joseph Antoine Ferdinand Plateau, 1801–1883. Belgium physicist.

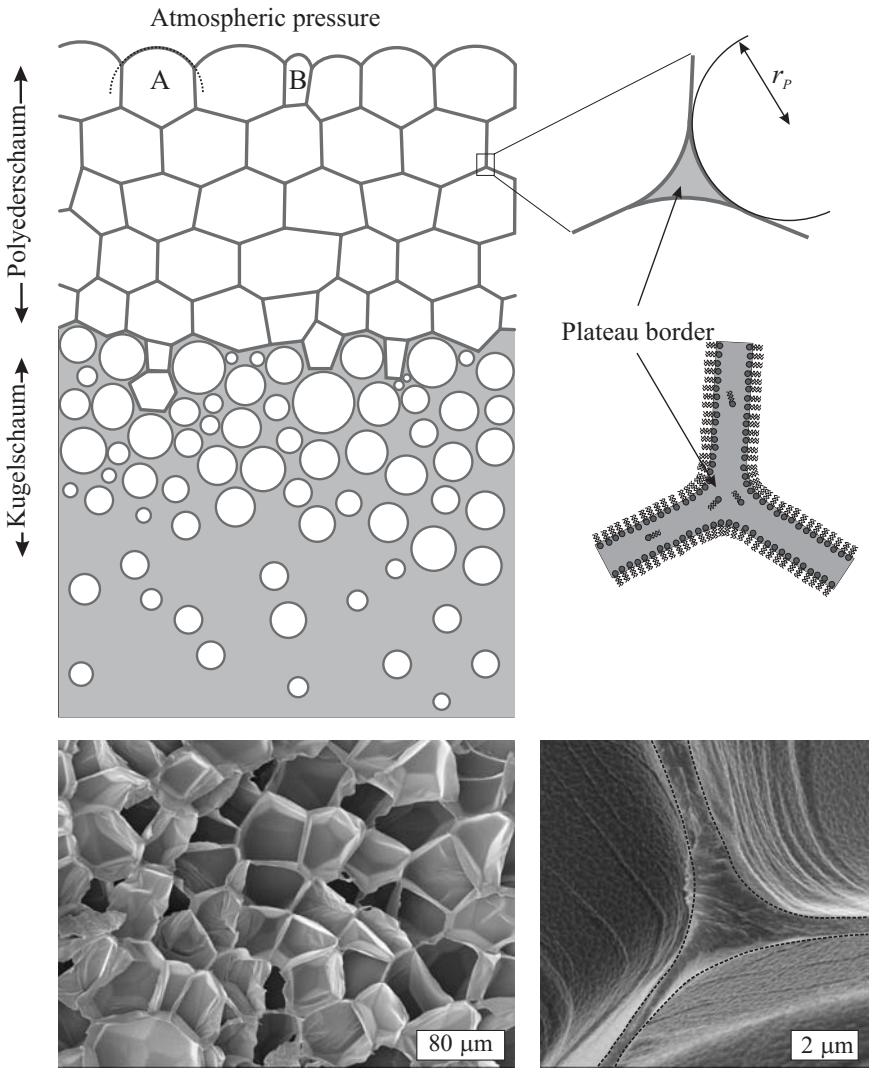



**Figure 12.18:** Top: Foam which forms on top of a liquid from rising bubbles. At the bottom of the froth the bubbles are densely packed and form a Kugelschaum. When the liquid is sufficiently drained out of the intervening spaces by gravitation, a polyhedral foam is formed. The detail on the right shows a cross-section through a Plateau border. The radius $r_P$ determines the reduced pressure in the Plateau border which is relevant for driving the liquid out of the planar films. Bottom: Scanning electron microscope image of a polystyrene foam. On the bottom right a cross-section through a Plateau border is shown. To expose the lamellae the foam was broken. This also explains the scratches in the Plateau border.

cell and thus to the film. The pressure inside the liquid film, which is equal to the disjoining pressure, is

$$\Pi = P_g - P_r + \frac{2\gamma}{r_C} - \rho g h \tag{12.24}$$

Here, $r_C$ is the inner radius of the capillary, $h$ is the height of the liquid column, and $\rho$ is the density of the liquid. We assume that the liquid wets the inner surface of the capillary. The pressure inside the liquid can be varied by changing $P_g$ with a pump. To measure the thickness of the film (or the distance between the two liquid–gas interfaces) white light is focussed from a normal direction onto the film. The light is reflected from both sides of the film and it interferes. The intensity of the reflected light is measured. From the interference the thickness can be calculated.

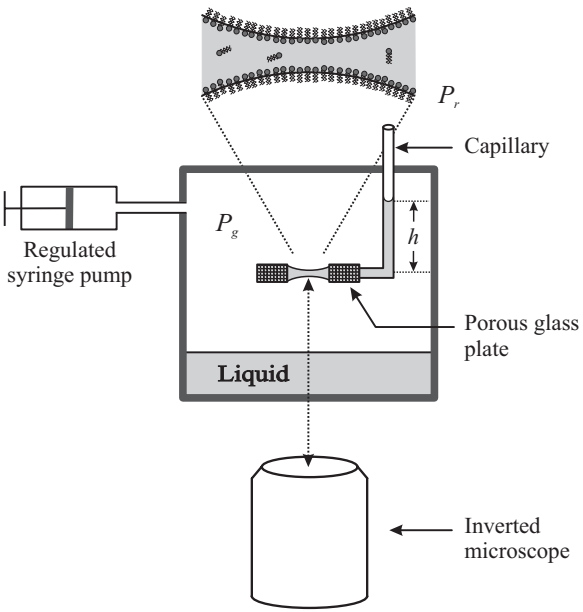



**Figure 12.19:** Schematic of a thin film balance.

When looking at soap films they appear look very colorful. These colors can be observed in reflection and transmission. They originate from the same effect as colors displayed by thin oil films on water, interference of light being reflected at the two interfaces. For this reason they critically depend on the film thickness and people who are experienced in dealing with soap films can tell the film thickness, just from looking at the colors of the film. It is not trivial to relate the color observed to the film thickness when illuminating with white light, because we not only have to consider a whole range of wavelengths, but we also have to take into account multiple reflections. A table of colors which were observed, related to film thicknesses and a description of the mathematical treatment is given in Ref. [565].

One case, that of a black film, is particularly important and easy to understand. Soap films, which are thinner than $\approx 30$ nm appear black. For this reason they are called black films. Why

are they black? When looking at a soap film at a not too large angle $\vartheta$ (Fig. 12.20) light being reflected at point A and at point B falls onto the eye of the observer. This light interferes and we only observe the resulting intensity. The path difference between the two rays corresponds to the distance ABC (since the distance between points A and C is small compared to the distance to the observer as the two reflected light rays are practically parallel). Since a black film is very thin the distance ABC is negligibly small compared to the wavelength of visible light $\lambda$. Light being reflected at the back side (point B), however, is phase shifted by $\lambda/2$ because it comes from the optically dense medium. For this reason the phase difference between the two rays is $\approx \lambda/2$ and they interfere destructively. No light is reflected and the film appears black.

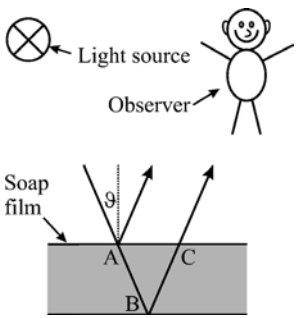



**Figure 12.20:** Schematic drawing of reflected interfering light rays from the two surfaces of a soap film.

### 12.5.4 Evolution of foams

Liquid foams evolve mainly by two processes: coarsening and drainage. The first process is slow and becomes effective after many minutes, hours or even days. Drainage under the influence of gravitation is faster and it is usually the main process, which destabilizes foams.

Let us start with coarsening. The compartments in a foam show a distribution of sizes. The smaller the compartment, the larger the surface-to-volume ratio. The whole system can decrease its Gibbs free energy by transporting gas from small to large compartments, which finally results in a coarsening also called Ostwald ripening: The large bubbles grow at the expense of small bubbles. The small bubbles will disappear. This process is possible because gas molecules can diffuse through the thin liquid films.

To gain more insight into the stability of foams let us have a closer look at the pressures involved. We refer to Fig. 12.18. Outside the foam the pressure is fixed. Usually this will be atmospheric pressure. The pressure in one of the top compartments is determined by the radius of curvature of the liquid film according to the Laplace equation. If we take water as a liquid with surfactant present and a typical surface tension of 40 mN/m the pressure inside one of the top compartments with a radius of curvature $R$ = 1 mm is $\Delta P = 4\gamma/R \approx 160$ Pa. The factor 4 is due to the fact that we do not have a bubble (then we would have $\Delta P = 2\gamma/R$) but a soap film with two sides. For compartment A the curvature is indicated by a semicircle.

Not all top compartments have precisely the same pressure inside. Some compartments are larger then others. In small compartments, for example, in compartment B, the pressure is higher. Then, gas will diffuse through the liquid films to neighboring compartments leading to coarsening. Large compartments grow in size while small compartments shrink.

To understand drainage we have to discuss the pressure inside the liquid films. At the contact line between liquid films, a channel is formed. This is called the Plateau border. Due to the small bending radius ($r_P$ in Fig. 12.18), there is a significant Laplace pressure difference between the inside of the compartment and the liquid phase. The pressure inside the liquid is significantly smaller than in the gas phase. As a result, liquid is sucked from the planar films into the Plateau's border. This is an important effect for the drainage of foams because the Plateau borders act as channels. Hydrodynamic flow in the planar films is a slow process [574]. It is for this reason that viscosity has a drastic influence on the evolution of a foam. Once the liquid has reached a Plateau border the flow becomes much more efficient. The liquid then flows downwards driven by gravitation.

In many applications we want to prevent the formation of foam or get rid of already existing foam (see, for example, Table 11.1). Usually chemicals are added to achieve this (see Ref. [564] for an introduction). We distinguish between antifoamers and defoamers. Antifoamers are added to the liquid prior to foam formation and act to prevent or inhibit foam formation. Defoamers or foam breakers are added to eliminate existing foam. They can only reach the outer surface of a foam.

Several chemicals destabilize foams by reducing the viscosity of the liquid and hence hasten drainage. Another way of destabilizing foams is to take surfactants away from the surface. Ionic surfactants can, for instance, be bound by oppositely charged inorganic species, which then form insoluble complexes. Undissolved oil droplets can spread at the interface and lead to rupture of soap films. In some cases hydrophobic particles are efficient defoamers for aqueous foams. The effect of particles is, however, complex and they might also stabilize a foam. Often, the addition of chemicals is not desired because they might contaminate the product or might cause pollution. Then a possible alternative method is to apply weak shock waves or mechanical vibrations to destroy the foam.

## 12.6 Summary

- We distinguish four types of surfactant. Anionic (e.g. SDS), cationic (e.g. CTAB), nonionic (e.g. alkylethylene glycols), and zwitterionic (e.g. phosphatidyl choline). Beside conventional surfactants, Gemini, bolafarm, oligomeric and polymeric surfactants become more and more important.

- In water, surfactants form spontaneously defined aggregates such as spherical micelles, cylinders, or bilayers, once the concentration has exceeded the CMC. Which aggregate is formed is largely determined by the surfactant parameter.

- A typical spherical micelle contains 30–100 surfactants and has diameter of 3–6 nm. Micelles form because of two competing factors: transfer of hydrocarbon chains out of water into an oil-like interior and repulsion between the head groups.

- Emulsions are dispersions of two immiscible liquids, usually water and oil. We distinguish macro- and microemulsions. Macroemulsions are not thermodynamically stable. Droplets are typically $0.5 - 10\mu m$ in diameter. Microemulsions are thermodynamically stable and droplets are of the order of 5–100 nm in diameter.

- Whether an oil-in-water or a water-in-oil emulsion is formed is largely determined by the surfactant. For macroemulsions Bancroft's rule and, more quantitatively, the concept of the HLB, are useful. They reflect the solubility of the surfactant in water and oil.
- The properties of ionic surfactants are mainly influenced by the salt concentration. For nonioinic surfactants the temperature is the most important parameter. Nonionic surfactants tend to favor oil-in-water emulsions below the PIT. Above the PIT water-in-oil emulsions are preferred.
- Foams are always thermodynamically stable. The stability of liquid foams is largely determined by the repulsion between surfactants and the viscosity of the liquid. They decay by drainage driven by the negative Laplace pressure in the Plateau borders.

## 12.7 Exercises

1. Micellization is driven by the hydrophobic effect. Estimate from measured CMCs of alkylethylene glycols, the change in the Gibbs free energy for bringing one methylene group ($\sim CH_2 \sim$) from an aqueous medium into the interior of a micelle.
2. Derive the equation for the volume of a hydrocarbon chain using the densities of $n$-hexane ($C_6H_{14}$, $\rho = 654.8$kg/m$^3$, $M = 86.18$ g/mol) and $n$-dodecane ($C_{12}H_{26}$, $\rho = 748.7$ kg/m$^3$, $M = 170.34$ g/mol).
3. Derive Eq. (12.13).
4. How much does the surface diminish if two spherical droplets with radius $r$, merge?
5. Flocculation kinetics. We have droplets in an emulsion with a density of one droplet per (10 $\mu$m)$^3$. Calculate the initial decrease of the concentration, assuming no energy barrier.
6. Elastic energy of a surfactant film. Please estimate the bending energy per unit area for a surfactant film with a bending rigidity of $10k_BT$ and zero spontaneous curvature, which is at the interface of a drop of radius 5, 20, and 100 nm.

# 13 Thin films on surfaces of liquids

## 13.1 Introduction

The adsorption of amphiphilic molecules at the surface of a liquid can be so strong that a compact monomolecular film, abbreviated as monolayer, is formed. There are amphiphiles which, practically, do not dissolve in the liquid. This leads to insoluble monolayers. In this case the surface excess $\Gamma$ is equal to the added amount of material divided by the surface area. Examples of monolayer forming amphiphiles are fatty acids ($CH_3(CH_2)_{n_C-2}COOH$) and long chain alcohols ($CH_3(CH_2)_{n_C-1}OH$) (see section 12.1).

An important class of practically insoluble amphiphiles are phospholipids (Fig. 13.1), which are an essential component of biomembranes. Phospholipids are 1,2-diesters of fatty acids and glycerol. At the third carbon atom the glycerol is further esterified to a phosphoryl ethanolamine, choline, or another polar group. There are different classes of lipids which differ in the chemical composition of their headgroups. Phosphatidyl cholines (PC) are zwitterionic in water because at neutral pH the phosphate is negatively charged ($H^+$ is dissociated) and choline is positively charged (it binds $H^+$). Phosphatidyl ethanolamines (PE) are another example of zwitterionic amphiphiles. Sugars as rest groups are uncharged so that, due to the negatively charged phosphate, the whole molecule is negative. An example is the phosphatidyl glycerols (PG). Phosphatidyl serins (PS) are also negatively charged, since the rest group is zwitterionic.

Within each class the phospholipids can have alkyl chains of different length (Table 13.1) and different numbers of double bonds. Alkyl chains which only consist of single bonds are called "saturated". Alkyl chains which contain at least one double bond are called "unsaturated". Unsaturated alkyl chains are more flexible and have lower melting temperatures.

| $n_C$ | Prefix | Common abbreviation |
|---|---|---|
| 12 | Dilauroyl- | DLPC / DLPE |
| 14 | Dimyristoyl- | DMPC / DMPE |
| 16 | Dipalmitoyl- | DPPC / DPPE |
| 18 | Distearoyl- | DSPC / DSPE |

**Table 13.1:** Nomenclature of dialkylglycerophosphatidyl choline (PC) and dialkylglycerophosphatidylethanolamine (PE) with alkyl chains of similar length.

If there is still a significant proportion of the amphiphile dissolved in the liquid we talk about Gibbs monolayers. Solubility in water is increased by using molecules with short alkyl chain or a high polarity of the headgroup. In this case $\Gamma$ is determined from the reduction of the surface tension according to the Gibbs adsorption isotherm (Eq. 3.52).

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

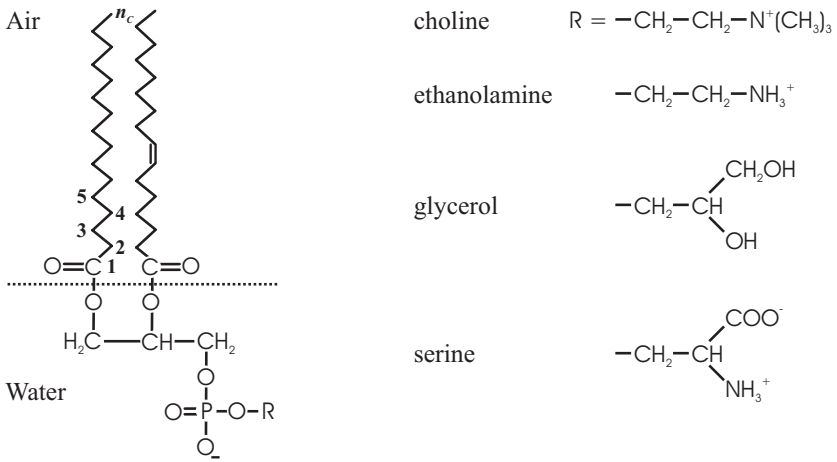



**Figure 13.1:** Structure of a phospholipid. Phopholipids can have different polar rest groups (R). In addition, the alkyl chains can have different lengths and can be saturated (left chain) or unsaturated (right chain).

The most important tool for studying insoluble monolayers is a film balance [575, 576], also called a Langmuir trough (Fig. 13.2). The modern version of a film balance consists of a temperature-controlled trough which contains the liquid. The liquid is called "subphase". Usually water is used as the subphase. The lipids are dissolved in a solvent (often chloroform) which is volatile and not miscible with the subphase. Drops of the solution are placed on the liquid surface and after evaporation of the solvent a lipid film remains. This process is called "spreading". Via movable barriers the film balance allows to manipulate the density of molecules on the liquid surface by compression or expansion of the film. When compressing the film the area per molecule decreases, when expanding the film, it increases.

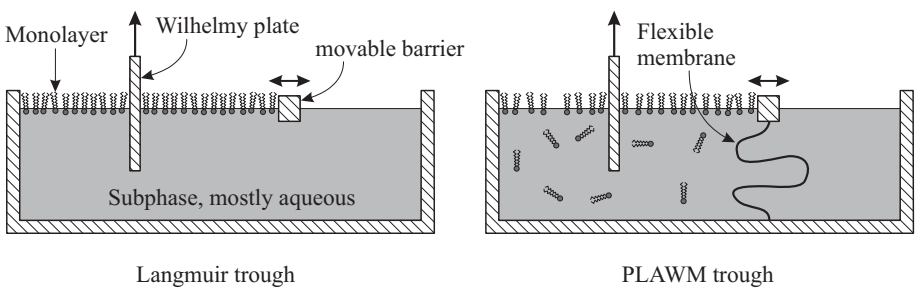



**Figure 13.2:** Langmuir and PLAWM trough with a monolayer indicated by amphiphilic molecules.

If the barrier could move freely, it would drift in the direction of the liquid with higher surface tension. In this way the system can reduce its entire free energy. We can imagine that this movement is caused by a film pressure, also called "lateral pressure". The film pressure

$\pi$ is defined as the difference between the surface tension of the bare subphase $\gamma_0$ and the surface tension of the subphase covered by amphiphiles $\gamma$:

$$\pi \equiv \gamma_0 - \gamma \tag{13.1}$$

The film pressure is usually measured by the Wilhelmy plate method. Usually the Wilhelmy plate is a piece of absorbent paper hanging into the water subphase. The force acting on it is proportional to the surface tension. More rarely, the force on the barrier is determined directly.

If we compress a surfactant film on water we observe that the surface tension decreases and the surface pressure increases. What is the reason for this decrease in surface tension? We can explain it by use of the Gibbs adsorption isotherm (Eq. (3.52)). On compression, the surface excess increases and hence the surface tension has to decrease. This, however, is relatively abstract. A more illustrative explanation is that the surface tension decreases because the highly polar water surface (high surface tension) is more and more converted into a nonpolar hydrocarbon surface (low surface tension).

If the amphiphile is significantly soluble in the liquid, we cannot use a Langmuir trough any longer because amphiphiles would diffuse via the liquid phase to both sides of the barrier. We can use a PLAWM[1] trough instead [577, 578]. In the PLAWM trough a flexible membrane, which is fixed to the barrier, separates the two compartments. This membrane is easily movable, so that the barrier position is affected only by surface effects.

At low surface excess, Gibbs monolayers can often be described as two-dimensional gases. This description is based on the observation that, at low concentration, the surface tension decreases linearly with the concentration of the added amphiphile $c$:

$$\gamma = \gamma_0 - bc \tag{13.2}$$

Here, $b$ is a constant, which depends on the solvent and the amphiphile. We insert this expression into the Gibbs adsorption isotherm (Eq. (3.52)):

$$\Gamma = -\frac{c}{RT}\frac{d\gamma}{dc} = \frac{bc}{RT} \tag{13.3}$$

Inserting Eq. (13.2) into the definition of the film pressure, Eq. (13.1), we immediately get $\pi = bc$. With this

$$\Gamma = \frac{\pi}{RT} \Rightarrow \pi = \Gamma RT \tag{13.4}$$

Substituting the surface excess by the inverse of the area per molecule, $\Gamma = 1/\sigma_A$, we get $\pi\sigma_A = RT$ or

$$\pi\sigma_A = k_B T \tag{13.5}$$

In the first equation $\sigma_A$ is in units of area per mole, in the second it is given in area per molecule.

[1] The name is composed of the first letters of Agnes Pockels (1862–1935, lady amateur scientist from Braunschweig, Lower Saxony), Langmuir, Adam, Wilson, and McBain.

At this point we need to be careful with the use of the symbol $\sigma_A$. Here, it is the average area available for one molecule and defined to be the inverse of the surface excess. In chapter 9 it is the related to the geometric size of a molecule. For example, if a molecule has a diameter of say 1 nm and it is adsorbed to a surface, its contact area is roughly $(1\text{nm})^2$. This is independent of the surface excess, except for very high surface coverage.

■ **Example 13.1.** What is the film pressure of sodium dodecylsulfate(SDS) at 25°C after adding 0.5 mM to water? From example 3.3 we know that, at this concentration, the area available to one molecule is 1.42 $\text{nm}^2$. This leads to a film pressure of

$$\pi = \frac{k_B T}{\sigma_A} = \frac{4.12 \times 10^{-21}\,\text{J}}{1.42 \times 10^{-18}\,\text{m}^2} = 2.90 \times 10^{-3}\,\frac{\text{N}}{\text{m}} \tag{13.6}$$

Instead of the ideal two-dimensional gas equation, often a van der Waals type equation is used, which contains an excluding surface $\sigma_0$:

$$\left(\pi + \frac{a}{\sigma_A^2}\right) \cdot (\sigma_A - \sigma_0) = k_B T \tag{13.7}$$

Here, $a$ is a material-dependent constant. Basic lectures on physical chemistry usually start by showing how the three-dimensional van der Waals equation of state explains condensation. Therefore the pressure is plotted versus the volume at constant temperature. By analogy, when plotting the film pressure $\pi$ versus the molecular area $\sigma_A$ at constant temperature, a two-dimensional condensation of monolayers is predicted, at least for temperatures below a certain critical temperature.

## 13.2 Phases of monomolecular films

The analogy between three- and two-dimensional phase diagrams can be carried much further. Monomolecular amphiphilic films show ordered phases similar to three-dimensional systems [579]. The phases of an amphiphilic monolayer can be detected most conveniently in pressure-area ($\pi$-versus-$\sigma_A$) isotherms. These may look different for different substances. The behavior of simple amphiphilic molecules, like long-chain alcohols, amines, or acids, was extensively investigated (reviews: Refs. [580, 581]). In monolayers so-called mesophases can occur. In a mesophase the tail groups are ordered over relatively large areas, while the order in the hydrophilic head groups is only over a much smaller distances.

Fatty acids, phospholipids, etc., on water often show the following phases [582–584] (Fig. 13.3):

- **Gaseous (G)**. For very large molecular areas, films are in the gaseous state. The gaseous state can be described by Eq. (13.7). The film pressure is, however, usually so small that it is almost undetectable. The average area per molecule on the surface is much larger than the size of the molecule. The total area can be expanded indefinitely without having a phase transition.

- **Liquid (L)**. When compressing a gaseous film, there can be a first-order phase transition to the liquid state. The liquid state is characterized by a significant lateral interaction

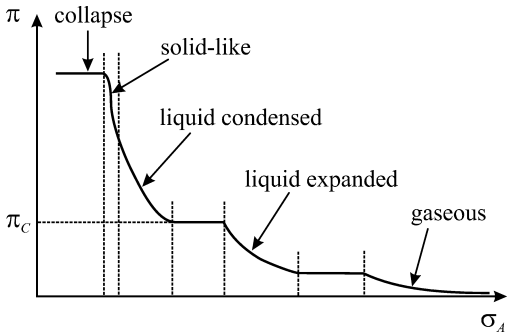



**Figure 13.3:** Schematic graph of film pressure versus surface area per molecule. The indicated phases can occur, but most amphiphiles do not show all phases.

between the amphiphiles. At least two types of liquid phase exist: a liquid-expanded (LE or $L_1$) and a liquid-condensed (LC or $L_2$) phase. For the liquid-expanded phase, pressure–area graphs extrapolated to a zero film pressure, show $\sigma_A$ values larger than the actual size of the molecule. For long-chain hydrocarbons with a polar head (e.g. fatty acids) this area is typically $\sigma_A \approx 40 - 70$ Å$^2$. The molecules are touching each other but there is no lateral order. The head groups are highly hydrated. The equation of state is often of van der Waals type.

After passing a plateau at a critical film pressure $\pi_c$ the liquid-condensed phase is reached via a phase transition of first order. Here, the amphiphiles exhibit a tilted phase with a decreasing tilt angle (measured against the normal to the subphase). The film is relatively stiff but there is still some water present between the headgroups.

- **Solid (S)**. The headgroups are largely dehydrated. Pressure–area isotherms are linear. Extrapolation to zero film pressure results in an area per molecule that corresponds to the molecular cross-section. For example, lipids with two long-chain fatty acids occupy an extrapolated area of $\approx 41$ Å$^2$, which corresponds to the cross-sectional area of the molecule [585].

Figure 13.4 shows three different pressure–area isotherms measured by DPPC at different temperatures. Real isotherms deviate in two aspects from the ideal, schematic ones. First, the phase transition between the liquid expanded and liquid condensed phase at $\sigma_A = 50...52$ mN/m is not sharp but smooth. This is due to the fact that the liquid condensed phase is still relatively soft and can still easily be compressed. Thus, the pressure increase during the phase transition not only causes more molecules to condense but also compresses the already formed liquid condensed phase. In addition, contaminations contribute to this effect.

Contaminations are also responsible for the second difference between real and ideal isotherms. At $\pi_c$ the isotherm is not perfectly horizontal but slightly tilted, in particular at elevated temperatures. Contaminations are expelled from the liquid condensed phase. Thus, when more and more of the monolayer goes into the liquid condensed phase, contaminations are enriched in the remaining liquid expanded phase. This reduces the two-dimensional

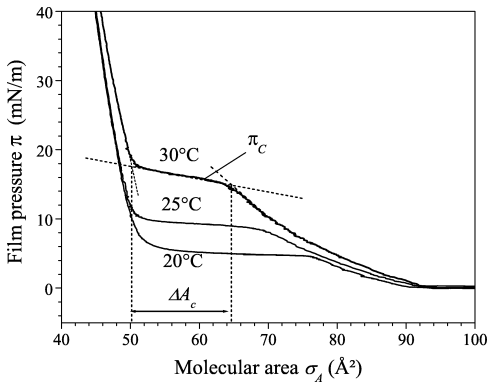



**Figure 13.4:** Pressure–area ($\pi$-versus-$\sigma_A$) isotherms recorded at different temperatures with DPPC [586]. For the isotherm recorded at 30°C the transition pressure $\pi_C$ is indicated.

concentration of the amphiphiles in the liquid expanded phase. To reach the concentration required for condensation, a higher pressure has to be applied.

The phase behavior of monolayers is determined by the molecular structure of the amphiphile and the conditions of the subphase. Phospholipids, for example, attract each other because of van der Waals interactions between the alkyl chains. The longer the alkyl chains, the more strongly the phospholipids attract each other. Thus, the LE–LC transition pressure will decrease with increasing chain length (at constant temperature). Double bonds in the alkyl chains increase this phase transition pressure. Charges and oriented dipole moments (see Chapter 6) in the headgroups, lead to a repulsion between the phopholipids and increase the pressure at which the transition occurs. Salts in the subphase, screen this repulsion and decrease the transition pressure.

■ **Example 13.2.** It is instructive to relate the film pressure to a three-dimensional pressure (Fig. 13.5). On the barrier of length $l$ the film exerts a force $\pi l$. In the three-dimensional case we estimate the force from the pressure $P$ which acts upon a surface $ld$, where $d$ is the thickness of the monolayer. This force is $Pld$. If the forces are set equal, we obtain $P = \pi/d$. Typical values for a monolayer in the $L_1$ phase are $d$ =1 nm and $\pi = 10^{-3}$ N/m. Then we estimate a three-dimensional pressure of $P = 10^6\,\mathrm{N/m^2} = 10\,\mathrm{atm}$.

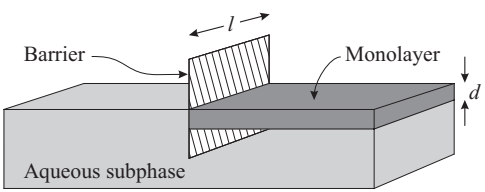



**Figure 13.5:** Monolayer on a water subphase compressed by a movable barrier.

Pressure–area isotherms of polymers or proteins often show no defined phases. Even at very low film pressures the behavior is not ideal. The behavior depends strongly on the specific structure of the polymer and the subphase, and isotherms are often irreversible. In general, the

molecular conformation at the air–subphase interface is a compromise between the polymer–polymer and polymer–subphase interactions on one hand, and the entropic contributions on the other hand.

## 13.3 Experimental techniques to study monolayers

The formation of monolayers and their thermodynamic investigation was described in the last chapter. A good introduction to the classic experimental techniques and results is given by Refs. [587] [588]. We start by discussing optical techniques.

### 13.3.1 Optical methods

In **fluorescence microscopy** (FM) a small amount of a fluorescent dye is added to the monolayer. To be incorporated into the monolayer the dye must be amphiphilic. The film is illuminated and the lateral distribution of the fluorescent molecules is observed with an optical microscope [589]. Depending on the phase condition of the monolayer, the fluorescent molecules distribute unevenly or have a different quantum yield. Usually the dyes are expelled from condensed liquid and solid phases. With this technique the coexistence of different phases in monolayers on water was demonstrated for the first time [590, 591].

■ **Example 13.3.** L-$\alpha$-DPPC (Dipalmitoylphosphatidylcholine) has a critical film pressure of $\pi_c$ = 12 mN/m on water at 22°C. The dye, in this case NBD-DPPC (NBD: 4-nitrobenzo-2-oxa-1,3-diazole), is expelled from the better ordered LC domains of DPPC (dark) and is enriched in the fluid LE-region (bright). The different phases can clearly be seen in the fluorescence image of Fig. 13.6. Upon compression, more and more of the dark appearing LC-domains will form until nearly all of the film becomes dark.

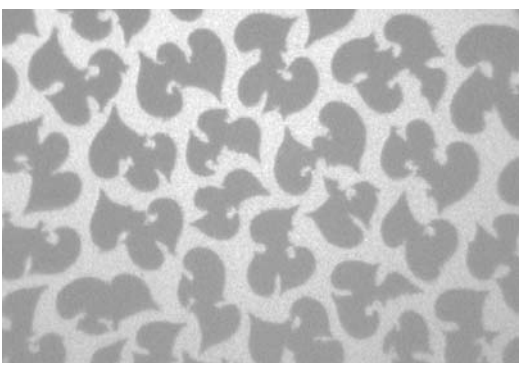


**Figure 13.6:** L-$\alpha$-DPPC containing 0.5% mol NDB-DPPC as a fluoroescent dye on water at 22°C at a film pressure of 12 mN/m. The picture (width 250 $\mu$m) was kindly provided by M. Lösche.

Example 13.3 demonstrates that phospholipids can form domains of distinct two-dimensional shapes on liquid surfaces. It has been found that the domain shape mainly depends on the chemical composition of the monolayer and the conditions such as temperature, pH, and ionic concentration. Domain structures can usually be understood by taking two competing interactions into account: an attractive dispersive van der Waals force and a repulsive dipole-dipole

interaction. The dipole–dipole interaction is repulsive because the molecules contain polar groups which are oriented parallel at the interface (see Fig. 13.7) and parallel dipoles repel each other. Moreover, the LC-domains form a perimeter against the surrounding LE-phase. This gives rise to a line tension which is based on the same physical reason as the surface tension: lipid molecules at the domain rim posses a higher energy than the lipids inside the domain. Thus, the formation of domains requires more energy, the longer is the perimeter for a given area. The electrostatic repulsion within the domain elongates it, the line tension tends to keep the perimeter circular. The actual shape is a trade-off between these two trends [580].

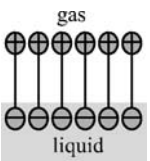



**Figure 13.7:** Schematic arrangement of dipolar amphiphiles in a monolayer.

An objection against fluorescence microscopy is that the presence of the fluorescent dyes possibly changes the structure of the monolayer. The problem is avoided in **Brewster angle microscopy** (BAM) [593–596]. In Brewster angle microscopy we use the following effect. Usually light shining on a liquid surface is partially reflected. The intensity of the reflected light depends on the angle of incidence. For light which is polarized parallel to the plane of incidence, there is one angle, under which no reflection occurs. This is the Brewster angle. For water, the Brewster angle is 53°. If we observe a water surface under an angle of 53° it appears dark. In Brewster angle microscopy the liquid is illuminated under the Brewster angle and its surface is observed using microscopy. The presence and structure of the monolayer slightly changes the Brewster angle. Regions which are covered with a monolayer have a slightly different Brewster angle and appear bright. With the help of Brewster angle microscopy the phase behavior of monolayers can be observed with a lateral resolution in the $\mu$m range [597]. Comparative studies with fluorescence and Brewster angle microscopy showed almost no differences [598].

### 13.3.2 X-ray reflection and diffraction

X-rays have a large penetration depth. In order to obtain a sufficiently intense signal from the surface layer the incident beam is applied under a very small angle (see exercise 8.4). Typical vertical irradiation angles $\alpha$ are 0.1°, which leads to penetration depths of $\approx 5$ nm. With a wavelength of a few Å (often the Cu–K$\alpha$ line with $\lambda = 1.54$ Å, is used) the X-rays are sensitive enough to analyze monolayers. Also thicker layers can be analyzed. Widely used X-ray techniques are X-ray reflection (XR) and diffraction (XD), which provide different information on thin films [585, 599, 601].

**X-ray reflection** ($\alpha' = \alpha, \beta = 0$). The intensity of the directly reflected beam is measured for different angles $\alpha$ (typically up to 5°). The experiments give information about the film thickness, the electron density distribution in electrons per Å$^3$ perpendicular (normal) to the liquid surface, and the roughness of the surface [602, 603]. To understand this we begin by discussing how film thickness is obtained.

If you have a thin film on top of a water surface, the incident X-ray beam is split into two beams; one being reflected from the film surface, the other being reflected from the film–water

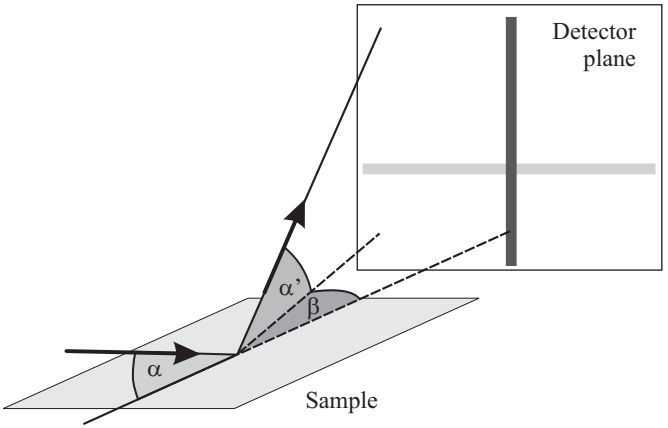



**Figure 13.8:** Set-up for X-ray reflection and diffraction experiments on monolayers.

interface (see Fig. 13.9). These beams interfere constructively or destructively (minimum) depending on the incident angle. The resulting reflected intensity shows maxima and minima. If we calculate the conditions for positive interference we get Bragg's law:

$$n\lambda = 2d \cdot \sin\alpha \tag{13.8}$$

Here, $\lambda$ is the wavelength of the X-rays, $n$ is the order of the minimum ($n = 1, 2, 3, \ldots$), and $d$ is the film thickness. Bragg's law implies that the higher the angle at which you find a minimum or maximum the smaller is the corresponding film thickness for a given wavelength.

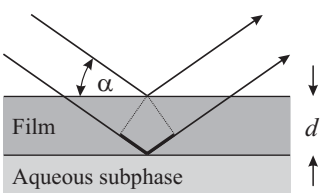



**Figure 13.9:** Interference of two X-rays reflected at the film surface and the film–liquid interface. The difference in path-length is highlighted. The angle of the incident X-rays, $\alpha$, in a real experiment is much lower.

X-rays interact with the electrons of the atoms in a material. Therefore, a necessary condition to resolve the additional film on the water surface is that the electron density of the film and the underlying liquid differ sufficiently. Additionally, for $\alpha' = \alpha$ there is no information on the horizontal component. Thus, the reflected intensity is an "electron density image" along the normal from air to water which is modified by a present monolayer film.

Real surfaces, especially those of liquids, are not perfectly flat. Even without external perturbations, thermal fluctuations in the form of capillary waves occur on liquid surfaces [602]. The higher the surface tension, the lower the resulting roughness. Usually, when compressing a monolayer to a high pressure (50 mN/m), i.e. low surface tension, the surface roughness of water changes from roughly 3 Å to 5 Å.

■ **Example 13.4.** The result of a typical X-ray measurement is shown in Fig. 13.10 for a galactocerebroside [605]. The plot on the left side shows the normalized reflected X-ray beam intensity versus the incident angle $\alpha$ for two different film pressures. The pressure–area isotherm is shown in the inset, together with the points of measurement a and b. On the right side are the extracted electron density profiles normal to the film surface taken at the same film pressures. At 0 Å we find the monolayer surface (top of the alkyl chains), a depth of –40 Å corresponds to pure water. In between is the film. The measurement is so sensitive that we even find two different electron densities within the monolayer. This is illustrated by the dashed boxes denoted by "film 1" and "film 2" (shown for curve b only) which represents the simplified electron density distribution in the so-called two-box model. A box is defined as a part in the film of a certain thickness where the electron density is constant. In the two-box model the film is divided into two layers. Film 1 represents the hydrocarbon tails, film 2 corresponds to the mean electron density of the head groups.

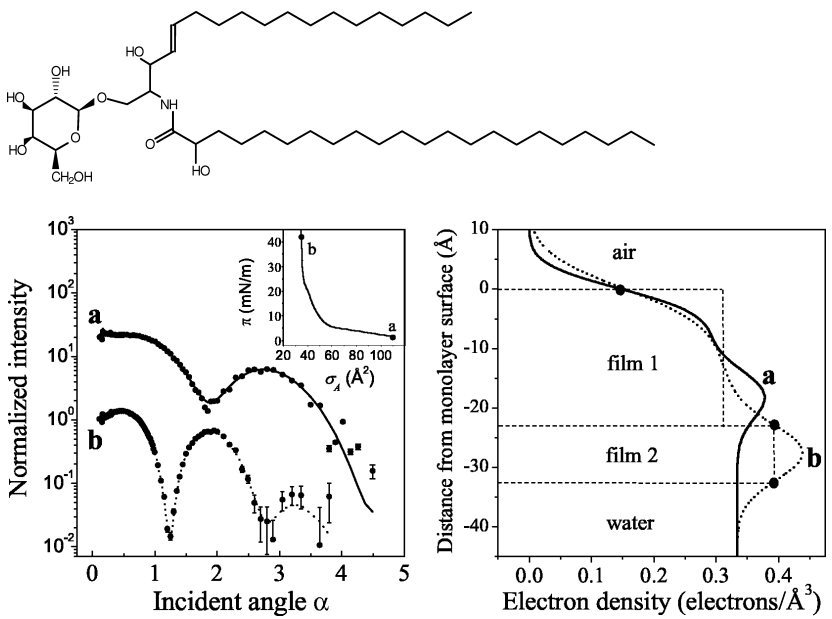



**Figure 13.10:** Structure of galactocerebroside (top) and result of an X-ray reflectometry experiment with galactocerebroside monolayer on water [605].

**X-ray diffraction** ($\beta \neq 0$). From the dependence of the in-plane reflexes from the angle $\beta$ we obtain information about the two-dimensional packing of the molecules on the surface

provided, they show some ordered structure. A perfect, flat two-dimensional crystal on the water surface would generate stripes ("Bragg rods") in the normal direction. If we analyze these rods as a function of the angle $\alpha$, information about the electron density profile parallel to the surface and about possible tilting angles of the molecules, is obtained. Diffraction measurements usually require intensive X-ray sources such as synchrotron radiation [607–610].

### 13.3.3 The surface potential

In general, monolayer-forming lipids contain polar groups which give rise to a dipole moment $\mu$. At the liquid surface all these dipoles are oriented parallel with a net dipole moment perpendicular to the subphase surface. Thus, a monolayer usually changes the electrical potential of the liquid–air interface. The change of the Volta potential between the liquid phase and the gaseous phase upon spreading of the layer is called the "surface potential" of the film. If $\chi_0$ is the potential between the inner part of the pure subphase and the gaseous phase and if $\chi$ is the potential in presence of the monolayer, then

$$\chi_{Surf} = \chi - \chi_0 \tag{13.9}$$

is the surface potential. $\chi_0$, $\chi$, and $\chi_{Surf}$ are counted as positive, if the electrode in the gaseous phase is more positive than the electrode in the liquid.

The surface potential is connected with the dipole moment of the molecules at the interface. If $\mu_\perp$ is the dipole moment normal to the surface, then the surface potential is given by [584]:

$$\chi_{Surf} = \frac{\mu_\perp}{\sigma_A \varepsilon \varepsilon_0} \tag{13.10}$$

Here, $\varepsilon$ is the dielectric constant of the liquid subphase. Practically, we measure the change of the surface potential when changing the film pressure. From this, we can get information on the dipole moment and orientation of the molecules on the surface [611–614].

In order to understand Eq. (13.10), we imagine that the dipoles form a homogeneous, infinitely extended, layer on the liquid surface (Fig. 13.11). The dipoles are supposed to have a dipole moment $\mu_\perp = Q \cdot d$. Directly on the liquid the monolayer starts with a charge density $\sigma = -\Gamma \cdot Q$. Then follows a region without free charges of thickness $d$. Finally we have a second layer with the charge density $\sigma = \Gamma \cdot Q$.

Now we try to find out which electric field strengths we have within the three regions. For simplicity we assume that there are no free charges in the liquid and we can set the electric field to zero throughout the liquid phase. In the region between the two charged layers the electric field strength is $E = \sigma/\varepsilon\varepsilon_0$. In the gaseous phase the field is zero again. To obtain the surface potential we integrate the electric field starting from somewhere in the liquid phase and ending somewhere in the gas phase. With this integration we get the potential:

$$\chi_{Surf} = \int_{liquid}^{gas} E dx = \int_0^d \frac{\sigma}{\varepsilon\varepsilon_0} dx = \frac{\sigma d}{\varepsilon\varepsilon_0} = \frac{\mu_\perp}{\sigma_A \varepsilon\varepsilon_0} \tag{13.11}$$

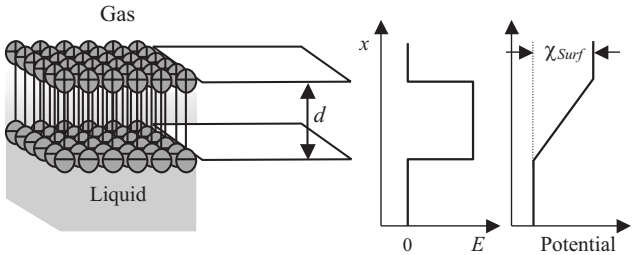



**Figure 13.11:** Schematic of the dipoles in a monolayer and the electric field strength $E$ and the potential versus the normal coordinate.

For molecules with no net free charge the measured dipole moment corresponds directly to the internal dipole moment of the molecule [615]. With charged molecules, an additional dipole has to be considered which is formed in the subphase due to the electrical double layer.

Measurements of the surface potential were first done in the 1932–37 [616, 617]. Two methods are commonly applied (fig. 13.12): the vibrating electrode and the ionizing electrode method. The **vibrating electrode method** goes back to Kelvin and therefore is also called a Kelvin probe. A plate-like electrode is attached at a distance $D$ above the liquid surface. It is connected to an electrode in solution. In order to measure the surface potential, the upper electrode is periodically moved up and down. This generates a current $I$. The generated alternating current is measured. The current is proportional to the surface potential according to [616]

$$I = \frac{dQ}{dt} = \chi_{Surf} \frac{dC}{dt} \tag{13.12}$$

From the capacity of a plate capacitor $C = \varepsilon_0 A/D$ it follows

$$\frac{dC}{dt} = \frac{dC}{dD} \cdot \frac{dD}{dt} = -\frac{\varepsilon_0 A}{D^2} \cdot \frac{dD}{dt} \tag{13.13}$$

Here, $A$ is the surface area of the plate capacitor. Substitution results in

$$I = -\chi_{Surf} \frac{\varepsilon_0 A}{D^2} \frac{dD}{dt} \tag{13.14}$$

■ **Example 13.5.** The electrode in air is 1 cm$^2$ large and 0.3 mm away from the surface of an aqueous electrolyte. It is periodically (in a sinusoidal way) moved up and down with an amplitude of $D_0 = 2.5$ $\mu$m and a frequency of $\nu = 330$ Hz. The used ampèrmeter can resolve a current $\Delta I = 1$ pA. How precisely can the surface potential be measured?

$$D = D_0 \cdot \sin(2\pi\nu t) \Rightarrow \frac{dD}{dt} = 2\pi\nu D_0 \cdot \cos(2\pi\nu t)$$

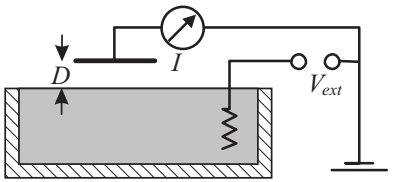



**Figure 13.12:** Schematic set-up to measure the surface potential of a monolayer.

The maximum difference in $dD/dt$ is $4\pi\nu D_0$. Thus

$$\begin{aligned}\Delta\chi_{Surf} &= \frac{\Delta I \cdot D^2}{4\pi\nu D_0 \varepsilon_0 A} \\ &= \frac{10^{-12}\,\mathrm{A}\cdot(3\times10^{-4}\,\mathrm{m})^2}{4\pi\cdot 330\,\mathrm{s}^{-1}\cdot 2.5\times10^{-6}\,\mathrm{m}\cdot 8.85\times10^{-12}\,\mathrm{AsV}^{-1}\,\mathrm{m}^{-1}\cdot 10^{-4}\,\mathrm{m}^2} \\ &= 0.01\,\mathrm{V}\end{aligned}$$

In the **ionizing electrode method**, the potential between a reference electrode in liquid (e.g. an AgCl electrode) and an upper electrode, which is placed few micrometers over the liquid, is measured. Practically, this is done by measuring the current as with the vibrating electrode method and applying an external voltage $V_{ext}$, which compensates the surface potential. When the external voltage just compensates the surface potential, the current $I$ is zero. In order to increase the conductivity of the air gap, the upper electrode is coated with an $\alpha$-emitter such as polonium $Po^{210}$.

### 13.3.4 Surface elasticity and viscosity

The mechanical characteristics of thin films on liquids are described in a similar way to the three-dimensional case. The surface elasticity [618] is defined as

$$E^\sigma \equiv A\cdot\frac{d\pi}{dA} = -A\cdot\frac{d\gamma}{dA} \tag{13.15}$$

Here, $A$ is the total surface area. $1/E^\sigma$ is also called compressibility. It represents a measure for the resistance of the film against compression. When measuring the surface elasticity we have to keep in mind that there may be different phases in the film which may have different mechanical properties [619].

The surface viscosity $\eta^\sigma$ is defined as follows. Two parallel line elements are moved parallel to each other. Between them (if we consider Newtonian behaviour) a constant speed gradient $dv/dy$ will exist. The force required to maintain the movement is

$$F = \eta^\sigma\cdot l\cdot\frac{dv}{dy} \tag{13.16}$$

where $l$ is the length of the line elements. Surface viscosity can be measured by a canal-type viscosimeter. In a canal viscosimeter we measure the surface area which flows at a certain film pressure per unit time through a gap or through a two-dimensional channel. With a surface

torsion pendulum the visco-elastic characteristics can be measured as well. In such a device we determine how quickly the rotational oscillation of a disk lying on the surface is damped. Usually, damping in pure liquid is always compared with the damping in the monolayer.

## 13.4 Langmuir–Blodgett transfer

Monolayers can be transferred layer-by-layer onto solid substrates such as silicon wafers or mica by the Langmuir–Blodgett[2] (LB) technique [620, 621] (reviews: Refs. [622–625]). This is shown schematically in Fig. 13.13.

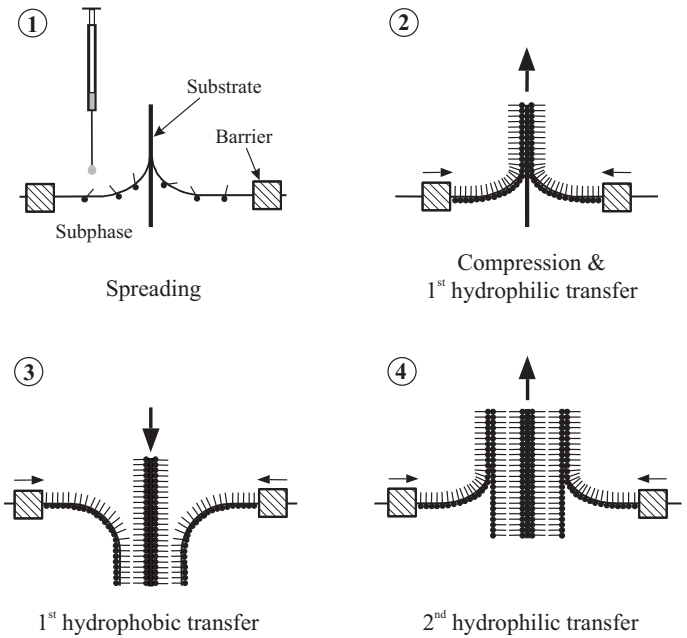



**Figure 13.13:** Langmuir–Blodgett transfer of organic monomolecular layers from water onto solid substrates.

The general procedure for a hydrophilic substrate is as follows:

1. The lipid is dissolved in a solvent which evaporates easily and is not miscible with water (usually chloroform, $CHCl_3$). After the hydrophilic solid substrate has been moved into the pure water subphase, drops of the lipid-containing solvent are set carefully onto the water surface between the movable barriers by a syringe ("spreading"). After solvent evaporation the monolayer is compressed to the desired pressure (usually some 20–40 mN/m, in the LC phase).

[2] Katherine Burr Blodgett, 1898–1979. American physicist.

2. The hydrophilic substrate is moved continuously out of the water subphase at constant film pressure. During the upstroke the monolayer is transferred onto the wafer with the headgroups oriented towards the solid substrate and the alkyl chains exposed to the air. This renders the hydrophilic solid surface with a high surface energy of about 50 mN/m (for silicon) to a hydrophobic surface with a relatively low surface energy in the range of 20–30 mN/m.

3. By a consecutive downstroke into the subphase through the floating monolayer a second layer can be transferred, with the alkyl chains oriented towards the solid substrate in a "tail-to-tail" configuration.

4. By another upstroke a third layer is transferred with the headgroups oriented towards each other ("head-to-head") and the hydrocarbon chains are exposed to the air.

Repeating these transfer cycles, multilayers can be formed on the substrate. The layers are called LB films. To investigate the structure of LB films the same techniques as above are usually applied, namely fluorescence microscopy, Brewster angle microscopy, and X-ray reflectometry.

The quality of a LB film, i.e. the ordering perpendicular to the substrate surface, depends on the quality of the first monolayer transfer. This is determined by the interaction between the substrate surface and the amphiphilic molecules and is usually influenced by the transfer film pressure and transfer speed [626–629]. This interaction can lead to a condensation of a lipid monolayer onto a silicon wafer before there is any condensed phase on the water surface. In a mixed monolayer film this condensation might lead to a demixing and thus separation of the single components during the film transfer. This process resembles a quasi two-dimensional zone melting, known from three-dimensional purification of silicon for use, e.g., in computer chips. There, a barrel of silicon containing impurities is moved slowly through a melting zone where the impurities are enriched in the melting part and can be removed after cooling [630].

The thickness of LB films can be of the order of the wavelength of visible light. Then they might appear colored or dark. This is due to the reflection and constructive or destructive interference of the light waves which are reflected from the film–air and the film–substrate surface. If the film has a thickness $d$ and a refractive index $n_1$, then for

$$d = \frac{\lambda}{4n_1} \tag{13.17}$$

the light interferes destructively and the film appears dark. Here, $\lambda$ is the wavelength of the incident light. This equation for the first-order minimum is valid for vertical light incidence and for $1 < n_1 < n_2$, with $n_2$ being the refractive index of the substrate. If we use a white light source with a whole range of different wavelengths, like sunlight, we would see the complementary colour of the destructively interfered wavelength. For example, if green is taken away by interference, the film appears red.

■ **Example 13.6.** Sodium salts such as NaCl exposed to a hot flame of a Bunsen burner emit bright yellow light of the wavelengths 589.0 nm and 589.6 nm (so-called Na D-lines) which is used as a hint for the presence of sodium in analytical chemistry. When we observe the reflection of this light from a LB film made of barium stearate ($n_1 = 1.491$) on a glass substrate ($n_2 \approx 1.52$) it appears completely dark for a film thickness of about 100 nm.

Nowadays, LB films are also used to cover electronical devices of special optical, electronical, or mechanical properties. Often, however, such hybrid systems between organic and inorganic compounds are not stable over a long time. For reviews on this topic see Refs. [631, 632].

## 13.5 Thick films – spreading of one liquid on another

What happens, when we put a drop of liquid B on the surface of liquid A? The two liquids are assumed not to mix and the density of B should be smaller than the density of A. That is, for example, the case for many oils on water. There are two possibilities: either liquid B spreads or it forms a lens with a defined edge.

Liquid B spreads spontaneously on liquid A, if the spreading coefficient, also called spreading pressure

$$S_{A/B} = \gamma_A - \gamma_B - \gamma_{AB} \tag{13.18}$$

is bigger than zero. Here, $\gamma_A$ and $\gamma_B$ are the surface tensions of the pure liquids and $\gamma_{AB}$ is the interfacial tension of the interface between the liquids. The condition results from a simple energy balance: For $S_{A/B} > 0$ it is energetically more favorable if liquid B spreads. More precisely, the Gibbs free energy of the system decreases if liquid B spreads and forms a homogeneous film (note: we are not talking about monomolecular films but thick films).

A complication now arises. $\gamma_A$ and $\gamma_B$ are values for the pure liquids. In a real situation, at least a tiny amount of molecules A is dissolved in liquid B, and vice versa. Therefore we should actually use $\gamma_{A(B)}$ and $\gamma_{B(A)}$ instead of $\gamma_A$ and $\gamma_B$. Here, $\gamma_{A(B)}$ is the surface tension of the liquid A, in which molecules B are dissolved up to saturation. For simplicity, we continue using the first notation.

If the spreading coefficient is smaller than zero, then a well defined drop of liquid B forms on the surface of A. The shapes of the surface of liquid A, the liquid B surface, and the liquid A–liquid B interface are determined by the Laplace equation. The angles $\Theta_1$, $\Theta_2$, and $\Theta_3$ are boundary conditions (Fig. 13.14). They are related by

$$\gamma_A \cos\Theta_3 = \gamma_B \cos\Theta_1 + \gamma_{AB} \cos\Theta_2 \tag{13.19}$$

The equation results from the following consideration. At equilibrium the three-phase contact line does not move. Hence, the sum of the forces acting in horizontal direction must be zero. The surface of liquid A pulls with a force (per unit length) of $\gamma_A \cdot \cos\Theta_3$ to the left. The surface of liquid B pulls to the right with a force $\gamma_B \cdot \cos\Theta_1$ and so does the liquid A–liquid B interface $\gamma_{AB} \cdot \cos\Theta_2$.

There is an interesting special case when liquid A has a much higher density than liquid B. This is, for example, the case if a water drop is put on mercury. In that case $\Theta_2 \approx 0$ and $\Theta_3 \approx 0$, and the boundary condition is simplified as:

$$\gamma_A = \gamma_B \cdot \cos\Theta_1 + \gamma_{AB} \tag{13.20}$$

The same equation applies to liquid drops on solid surfaces and leads to Young's equation as discussed in Chapter 7.

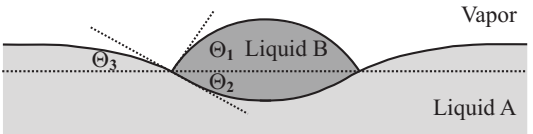



**Figure 13.14:** A drop of liquid B (e.g. oil) on the surface of an imissible liquid A (e.g. water).

How can we determine the interfacial tension between two imissible liquids $\gamma_{AB}$? The interfacial tension between two liquids can be measured by the spinning drop method (Fig. 13.15). Therefore, the liquid with the higher density – let it be liquid A – is filled into a horizontal capillary. Through a septum, which closes the capillary, a drop of liquid B is injected with a syringe into the center of the capillary. When the capillary is rotated rapidly liquid B will go to the center, forming a drop around the axis of revolution. With increasing speed of revolution the drop elongates, since the centrifugal force increasingly opposes the surface tensional drive toward minimum interfacial area. At sufficiently high speed of revolution, the drop is shaped like an elongated cylinder. From the shape we can calculate the interfacial tension. A summary of interfacial tensions between water and organic liquids is contained in Ref. [633].

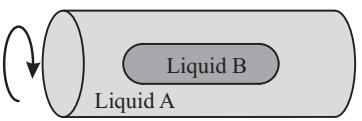



**Figure 13.15:** Spinning drop method to measure the interfacial tension between two liquids.

**Table 13.2:** Interfacial tensions $\gamma_{AB}$ of liquid–liquid interfaces at 20°C [1].

| Liquid A | Liquid B | $\gamma$ (mN/m) |
|---|---|---|
| Water | mercury | 375.0 |
| | $n$-hexane | 51.1 |
| | $n$-octane | 50.8 |
| | toluene | 36.1 |
| | benzene | 35.0 |
| | chloroform | 32.8 |
| | benzaldehyde | 15.5 |
| | $n$-octanol | 8.5 |
| | $n$-butanol | 1.8 |
| Mercury | ethanol | 389 |
| | $n$-hexane | 378 |
| Fluorocarbon polymer | benzene | 7.8 |

## 13.6 Summary

- Amphiphilic molecules form monomolecular films on liquid surfaces. Some amphiphiles, such as phospholipids, with a large hydrophobic tail are practically insoluble in water and therefore form insoluble monolayers at the air–water interface.

- A variety of methods were developed to investigate monolayers. In a film balance, also called a Langmuir trough, the film pressure versus the surface density or surface area is recorded. Electrical and rheological properties of the monolayer can be measured as the surface potential and the viscosity, respectively. Morphological changes within the monolayer can be observed by fluorescence microscopy or Brewster angle microscopy. The structural molecular ordering in monolayers is investigated by X-ray reflection and diffraction. X-ray reflection allows to determine the film thickness and the electron density distribution normal to the water surface. From this, details like surface roughness or hydration of the headgroups can be extracted. With X-ray diffraction, positional and oriental ordering within the monolayer are measured.

- Monolayers show different gas-, liquid-, and solid-like phases. This can be observed with a film balance. The shape of phase boundaries within monolayers are determined by the competition between electrostatic repulsion of the molecular dipole moments, the van der Waals attraction, and the perimeter-minimizing line tension.

- Monolayers can be transferred onto solid substrates by the Langmuir–Blodgett technique giving rise to mono- or multilayers which can even have thicknesses in the range of the wavelength of visible light. The multilayer quality often depends on the quality of the first monolayer transfer.

- When setting an immiscible liquid on top of another liquid, a macroscopic thick film or a drop forms, depending on the interfacial tensions. This is quantified using the spreading coefficient.

## 13.7 Exercises

1. A monolayer at the air–water interface is compressed to a pressure of 30 mN/m. Its thickness is measured as 25 Å. What is the corresponding three-dimensional pressure?

2. Isotherms of monolayers at the air–water interface are often measured at different temperatures $T$. It is found that they are usually shifted to higher critical pressures $\pi_c$ (of the $1^{st}$ order transition) with increasing $T$. A typical example is shown in Fig. 13.4. From the Clapeyron equation

$$\frac{\partial \pi_c}{\partial T} = -\frac{\Delta H_c}{T \Delta \sigma_c}$$

you can calculate the transition enthalpy $\Delta H_c$ and the transition entropy $\Delta S_c$ for the LE–LC phase transition. Determine $\Delta H_c$ and $\Delta S_c$ for each temperature.

3. A monolayer of hydroxygalactocerebroside ($C_{46}H_{89}NO_9$) is investigated by X-ray reflection on water (Fig. 13.10) [605]. The measured reflectivities ($I/I_0$, that is the reflected intensity divided by the initial intensity $I_0$) are plotted versus the angle of incidence $\alpha$ for two different film pressures $\pi_a$ and $\pi_b$ thus two different molecular areas. The wavelength was 1.54 Å.

   (a) Determine the corresponding film thicknesses for $\pi_a$ and $\pi_b$ from the first minimum in each reflectogram. Compare your result with the length as estimated from the molecular structure assuming completely stretched, all-trans alkyl chains.

   (b) What is the "effective" volume of each amphiphile? The "effective" volume includes the water molecules. Do you have an idea why they are different at different film pressures?

   (c) On the right side of Fig. 13.10 the two extracted electron density distributions with increasing depth are shown. Additionally, for film pressure b, the two-box model is shown. Determine the volume of a water molecule from the plot. Is your result reasonable?

   (d) The electron density distribution shows details of the film structure. What are the thicknesses $d_1$ and $d_2$? of the film substructures "film1" and "film2"? What are the corresponding electron densities $\rho_1$ and $\rho_2$? What do you think are the structures "film1" and "film2" if you compare your result with the molecular structure?

   (e) Determine the overall number of electrons $n_e$ in the whole molecule using the equation: $n_e = (d_1\rho_1 + d_2\rho_2)\cdot\sigma_A$. Compare your result with the real number of electrons from the molecular structure. Do you have an idea why there is a difference?

4. Derive Eq. (13.17). Note that the condition is $1 < n_1 < n_2$ and that there is an optical phase shift of $\pi$ if light is reflected from a material with higher optical density and a phase shift of 0 if the reflecting material has a lower optical density. What is the condition for destructive interference in the case $1 < n_1 > n_2$?

# 14 Solutions to exercises

## Chapter 2: Liquid surfaces

1. **Surface contamination**. For the estimation we use Eq. (2.1) and take nitrogen as a gas. With a molar mass of 28 g/mol the molecular mass is $m = \frac{0.028\,\mathrm{kgmol^{-1}}}{6.02\times10^{23}\,\mathrm{mol^{-1}}} = 4.7\times10^{-26}\,\mathrm{kg}$. The area occupied by one adsorbed molecule is of the order of $\sigma_A \approx 10\,\text{Å}^2 = 0.1\,\mathrm{nm}^2$. To get an average of 0.1 hits in $\sigma_A$ (this leads to 10% coverage) in a time span $\Delta t$ we solve

$$0.1 = \frac{P\sigma_A\Delta t}{\sqrt{2\pi m k_B T}} \Rightarrow$$

$$P = \frac{0.1\cdot\sqrt{2\pi\cdot4.7\times10^{-26}\,\mathrm{kg}\cdot1.38\times10^{-23}\cdot298\,\mathrm{J}}}{10^{-19}\,\mathrm{m}^2\cdot3600\,\mathrm{s}} = 9.7\times10^{-9}\,\mathrm{Pa}$$

We assumed that the coverage is low. Otherwise we had to solve a differential equation and obtain an exponential dependence on $\Delta t$.

2. **Sitting or pendent drop**. Both methods involve the determination of the shape of the drop in mechanical equilibrium. The shape is determined by the balance between gravitation and surface tensional forces. If gravitation is negligible the shape is always spherical irrespective of the surface tension.

3. **Water running out of a plastic box**. Since the plastic is not wetted by water, the maximum radius of a water drop coming out of the box is 100 $\mu$m. To push a water drop of 100 $\mu$m radius out of a hole, a pressure of

$$\Delta P = \frac{2\gamma}{R} = \frac{2\cdot0.072\,\mathrm{N/m}}{10^{-4}\,\mathrm{m}} = 1440\,\mathrm{Pa}$$

is required. This is equal to the hydrostatic pressure $P = \rho\gamma h$ at a height

$$\Delta P = \frac{2\gamma}{r} = \rho g h \Rightarrow h = \frac{2\cdot0.072\,\mathrm{N/m}}{100\times10^{-6}\,\mathrm{m}\cdot9.81\,\mathrm{m/s^2}\cdot1000\,\mathrm{kg/m^3}} = 0.147\,\mathrm{m}$$

Thus, the water runs out of the box until a height of 14.7 cm is reached.

4. **Wilhelmy-plate method**. The force on the plate is

$$2l\gamma\cos\Theta = 2\cdot0.01\,\mathrm{m}\cdot0.072\,\mathrm{N/m}\cdot0.707 = 1.02\times10^{-3}\,\mathrm{N}$$

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

5. **Drop-weight method**. With a mass $m = 2.2 \times 10^{-3}\,\mathrm{kg}/100 = 2.2 \times 10^{-5}\,\mathrm{kg}$ the volume of one drop is

$$V = \frac{m}{\rho} = \frac{2.2 \times 10^{-5}\,\mathrm{kg}}{773\,\mathrm{kg/m^3}} = 2.8 \times 10^{-8}\,\mathrm{m^3}$$

Since hexadecane wets the capillary we choose the outer diameter of the capillary and get

$$r_C/V^{1/3} = \frac{2 \times 10^{-3}\,\mathrm{m}}{(2.8 \times 10^{-8}\,\mathrm{m^3})^{1/3}} = 0.659$$

The correction factor is $f = 0.616$ as obtained by interpolation. Using this correction factor we get

$$\gamma = \frac{mg}{2\pi f r_C} = \frac{2.2 \times 10^{-5}\,\mathrm{kg} \cdot 9.81\,\mathrm{m/s^2}}{2\pi \cdot 0.616 \cdot 2 \times 10^{-3}\,\mathrm{m}} = 0.0279\,\frac{\mathrm{N}}{\mathrm{m}}$$

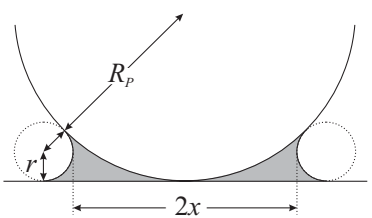
With the non-corrected equation we obtain 0.0172 N/m. Conclusion: The correction is very important! The real surface tension of hexadecane is 27.1 mN/m at 25°C.

6. **Capillary condensation**. We start with Kelvin's equation:

$$RT \cdot \ln \frac{P_0^K}{P_0} = \gamma V_m \cdot \left(\frac{1}{R_1} + \frac{1}{R_2}\right) = \gamma V_m \cdot \left(\frac{1}{x} - \frac{1}{r}\right) \approx -\frac{\gamma V_m}{r} \Rightarrow$$

$$r = -\frac{\gamma V_m}{RT \cdot \ln\left(P_0^K/P_0\right)}$$

using the same notation as in Section 2.6.

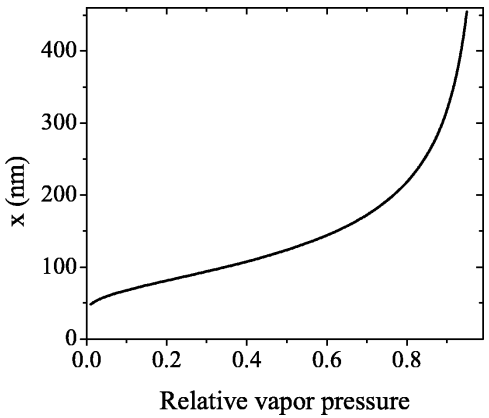



Roughly we have

$$R_P^2 + (x+r)^2 = (R_P + 2r)^2 \Rightarrow R_P^2 + x^2 + 2rx + r^2 = R_P^2 + 4rR_P + 4r^2$$
$$\Rightarrow x^2 + 2rx = 4rR_P + 3r^2 \Rightarrow x^2 \approx 4rR_P$$

Inserting

$$\begin{aligned} x &= \sqrt{-\frac{4\gamma V_m R_P}{RT \cdot \ln \frac{P_0^K}{P_0}}} = \sqrt{-\frac{4 \cdot 0.072 \frac{\mathrm{N}}{\mathrm{m}} \cdot 18 \times 10^{-6} \frac{\mathrm{m^3}}{\mathrm{mol}} \cdot 5 \times 10^{-6}\mathrm{m}}{8.31 \frac{\mathrm{J}}{\mathrm{mol\,K}} \cdot 298\,\mathrm{K} \cdot \ln \frac{P_0^K}{P_0}}} \\ &= 0.103 \cdot \mu\mathrm{m} \sqrt{-\frac{1}{\ln\left(P_0^K/P_0\right)}} \end{aligned}$$




7. **Force on a wire**. The force is given by

$$F = (2l + 4r)\,\gamma \Rightarrow \gamma = \frac{F}{2l + 4r} = \frac{0.49 \times 10^{-3}\,\mathrm{N}}{0.02\,\mathrm{m} + 0.004\,\mathrm{m}} = 20.4\,\frac{\mathrm{mN}}{\mathrm{m}}$$

For $0 < \Theta < 90°$ the wire is released at $F = (2l + 4r\cos\Theta)\,\gamma$.

## Chapter 3: Thermodynamics of surfaces

1. **Surface entropy and energy**. Methanol:

$$\begin{aligned} s^\sigma &= -\frac{\partial\gamma}{\partial T} \approx -\frac{\Delta\gamma}{\Delta T} = -\frac{\frac{0.02207\,\mathrm{N} - 0.02323\,\mathrm{N}}{15\,\mathrm{mK}} + \frac{0.02014\,\mathrm{N} - 0.02207\,\mathrm{N}}{25\,\mathrm{mK}}}{2} \\ &= \frac{7.73 \times 10^{-5}\,\frac{\mathrm{N}}{\mathrm{mK}} + 7.72 \times 10^{-5}\,\frac{\mathrm{N}}{\mathrm{mK}}}{2} = 3.87 \times 10^{-5}\,\frac{\mathrm{J}}{\mathrm{m}^2 K} \end{aligned}$$

$$u^\sigma = \gamma - T\frac{\partial\gamma}{\partial T} = 0.02207\,\frac{\mathrm{N}}{\mathrm{m}} + 298\,\mathrm{K} \cdot 3.87 \times 10^{-5}\,\frac{\mathrm{J}}{\mathrm{m}^2\,\mathrm{K}} = 0.0336\,\frac{\mathrm{N}}{\mathrm{m}}$$

Octane:

$$s^\sigma = 10.03 \times 10^{-5}\,\frac{\mathrm{J}}{\mathrm{m}^2\,\mathrm{K}} \qquad \text{and} \qquad u^\sigma = 0.0510\,\frac{\mathrm{N}}{\mathrm{m}}$$

2. **Pressure in bubble**. Without surfactant the pressure inside the bubble is:

$$\Delta P = \frac{2\gamma}{R} = \frac{2 \cdot 0.072\,\mathrm{N/m}}{0.01\,\mathrm{m}} = 14.4\,\mathrm{Pa}$$

With surfactant the surface tension is reduced according to the Gibbs adsorption isotherm Eq. (3.52). To apply Eq. (3.52) we need to know the surface excess:

$$\Gamma = \frac{1}{0.7 \times 10^{-18}\,\mathrm{m}^2} = 1.43 \times 10^{18}\mathrm{m}^{-2} = 2.38 \times 10^{-6}\,\frac{\mathrm{mol}}{\mathrm{m}^2}$$

$$\Gamma = -\frac{a}{RT} \cdot \frac{\partial \gamma}{\partial a} \approx -\frac{a}{RT} \cdot \frac{\Delta \gamma}{\Delta a} \Rightarrow \Delta \gamma = -\frac{\Gamma RT \Delta a}{a}$$

$$\Delta \gamma = -\frac{2.38 \times 10^{-6}\,\mathrm{mol/m^2} \cdot 2480\,\mathrm{J/mol} \cdot 0.002\,\mathrm{mol/L}}{0.002\,\mathrm{mol/L}} = 5.90 \times 10^{-3}\,\frac{\mathrm{J}}{\mathrm{m^2}}$$

The surface tension is reduced to $0.0720\,\mathrm{N/m} - 0.0059\,\mathrm{N/m} = 0.0661\,\mathrm{N/m}$ which leads to a pressure of 13.2 Pa.

3. **Adsorption from gas phase**. The concentration (activity) $a$ of gas molecules is proportional to the partial pressure $P$: $a = KP$, where $K$ is a constant. Inserting into Eq. (3.52):

$$\Gamma = -\frac{KP}{RT} \cdot \frac{\partial \gamma}{\partial (KP)} = -\frac{P}{RT} \cdot \frac{\partial \gamma}{\partial P} = -\frac{1}{RT} \cdot \frac{\partial \gamma}{\partial \ln P}$$

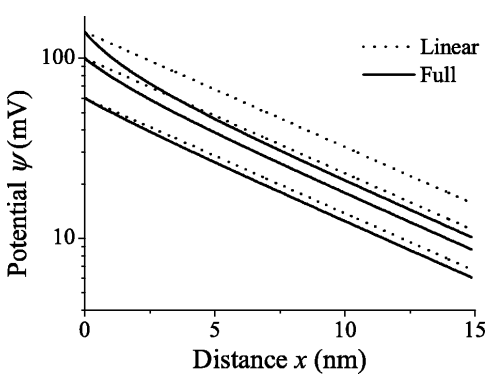
## Chapter 4: The electric double layer

1. **Debye length**. Inserting $\varepsilon = 25.3$ for ethanol and a concentration of $6.02 \times 10^{22}$ salt molecules per m$^3$ into Eq. (4.8), and with $\lambda_D = \kappa^{-1}$ we get

$$\lambda_D = \sqrt{\frac{25.3 \cdot 8.85 \times 10^{-12}\,\mathrm{AsV^{-1}\,m^{-1}} \cdot 1.38 \times 10^{-23}\,\mathrm{J\,K^{-1}} \cdot 298\,\mathrm{K}}{2 \cdot 6.02 \times 10^{22}\,\mathrm{m^{-3}} \cdot (1.60 \times 10^{-19}\,\mathrm{As})^2}} = 17.3\,\mathrm{nm}$$

For water the decay length is larger according to $\lambda_D = 0.304\,\mathrm{nm}/\sqrt{c_0} = 30.4$ nm.

2. **Potential-versus-distance at high surface potentials**. With a Debye length of 6.80 nm we get the following plot.




3. **Capacitance of an electric double layer**. From the concentration we calculate the Debye length: $c = 0.001\,\mathrm{M} \Rightarrow \lambda_D = 9.6\,\mathrm{nm} \Rightarrow C = 0.072\,\mathrm{F/m^2}$ and $c = 0.1\,\mathrm{M} \Rightarrow \lambda_D = 0.96\,\mathrm{nm} \Rightarrow C = 0.72\,\mathrm{F/m^2}$. To obtain the total capacities we have to multiply this with the surface area. Neglecting the end cap the surface area is $\pi \cdot 0.4 \times 10^{-3}\,\mathrm{m} \cdot 0.05\,\mathrm{m} = 6.3 \times 10^{-5}\,\mathrm{m^2}$. The total capacities are 4.5 $\mu$F and 45 $\mu$F, respectively.

4. **Cation concentration at a silicon oxide surface**. The concentration of cations is increased by a factor $e^{e\psi_0/k_BT} = e^{\frac{1.60\times10^{-19}\,\mathrm{As}\cdot0.07V}{4.12\times10^{-21}\,\mathrm{J}}} = 15.3$. Despite the low bulk concentration we therefore expect a concentration of 0.76 M. This corresponds to $4.6\times10^{26}$ counterions per $\mathrm{m}^3$. The average distance is only $\approx 1.3$ nm.

   The $H^+$ concentration at the surface is given by $[H^+] = 15.3\,[H^+]_0$, where $[H^+]_0 = 10^{-9}$ M is the bulk concentration of $H^+$. The local pH at the surface is $\log[H^+] = \log\left(1.53\times10^{-8}\right) = 7.8$.

5. **Capacitance of mercury**. For $\psi_0 = 0$ and considering that $\cosh 0 = 1$, Eq. (4.29) simply reads $C^A_{GC} = \varepsilon\varepsilon_0/\lambda_D$. With $\varepsilon = 78.4$ we get

$$\begin{aligned}
c_0 &= 0.001\,\mathrm{M} \Rightarrow \lambda_D = 9.61\,\mathrm{nm} \Rightarrow C^A_{GC} = 7.2\,\mu\mathrm{F/cm^2}\\
c_0 &= 0.01\,\mathrm{M} \Rightarrow \lambda_D = 3.04\,\mathrm{nm} \Rightarrow C^A_{GC} = 22.9\,\mu\mathrm{F/cm^2}\\
c_0 &= 0.1\,\mathrm{M} \Rightarrow \lambda_D = 0.961\,\mathrm{nm} \Rightarrow C^A_{GC} = 72.3\,\mu\mathrm{F/cm^2}\\
c_0 &= 1\,\mathrm{M} \Rightarrow \lambda_D = 0.304\,\mathrm{nm} \Rightarrow C^A_{GC} = 229\,\mu\mathrm{F/cm^2}
\end{aligned}$$

   Compared to the experimental results this clearly shows, that Gouy-Chapman theory fails to describe the capacitance of the double-layer, especially at high salt concentration.

## Chapter 5: Effects at charged interfaces

1. **Gibbs free energy of double layer around a spherical particle**. We can use an equation similar to Eq. (4.35), we only have to use the total charge of the particle instead of the surface charge density. The total charge of the particle is given by Eq. (5.39).

$$\begin{aligned}
G &= -\int_0^{\psi_0} Q d\psi_0' = -4\pi\varepsilon\varepsilon_0 R\left(1+\frac{R}{\lambda_D}\right)\cdot\int_0^{\psi_0}\psi_0' d\psi_0'\\
&= -4\pi\varepsilon\varepsilon_0 R\left(1+\frac{R}{\lambda_D}\right)\cdot\left[\frac{\psi_0'^2}{2}\right]_0^{\psi_0} = -\frac{Q^2}{8\pi\varepsilon\varepsilon_0 R\cdot\left(1+\frac{R}{\lambda_D}\right)}
\end{aligned}$$

2. **Electro-osmosis**. The flow velocity is

$$|v_0| = \varepsilon\varepsilon_0\frac{\zeta E_x}{\eta} = 78.5\cdot 8.85\times10^{-12}\frac{\mathrm{As}}{\mathrm{Vm}}\cdot\frac{0.03\,\mathrm{V}\cdot 100\,\mathrm{Vm^{-1}}}{0.001\,\mathrm{kgs^{-1}m^{-1}}} = 2.08\times10^{-6}\frac{\mathrm{m}}{\mathrm{s}}$$

   The flow is directed from plus to minus.

3. **Electrophoresis**. The drift velocity of the particles with respect to the liquid is given by Eq. (5.40). Inserting a Debye length of $\lambda_D = 0.304\,\mathrm{nm}/\sqrt{0.01} = 3.04\,\mathrm{nm}$ leads to

$$\begin{aligned}
|v_0| &= \frac{2\cdot 78.5\cdot 8.85\times10^{-12}\mathrm{AsV^{-1}m^{-1}}\cdot 0.02\mathrm{V}\cdot 100\mathrm{Vm^{-1}}}{3\cdot 0.001\,\mathrm{kgs^{-1}\,m^{-1}}}\cdot\left(1+\frac{50\,\mathrm{nm}}{3.04\,\mathrm{nm}}\right)\\
&= 0.92\frac{\mu\mathrm{m}}{\mathrm{s}}\cdot 17.4 = 16.1\,\frac{\mu\mathrm{m}}{\mathrm{s}}
\end{aligned}$$

   The particle drifts towards the positive electrode opposite to the liquid flow with a velocity of $16.1 - 2.1 = 14.0\,\mu\mathrm{m/s}$. It is certainly not a good idea to use them as markers.

## Chapter 6: Surface forces

1. **Van der Waals force of layer of molecules**. The energy between a single molecule A and a planar monolayer of surface density $\rho_B^\sigma$ is

$$\begin{aligned} W_{mol/pl} &= -C_{AB}\rho_B^\sigma \int_0^\infty \frac{2\pi r dr}{\left(D^2+r^2\right)^3} = -\pi C_{AB}\rho_B^\sigma \int_0^\infty \frac{dr^2}{\left(D^2+r^2\right)^3} \\ &= -\frac{\pi\rho_B^\sigma C_{AB}}{2D^4} \end{aligned}$$

The energy between for a whole planar layer of molecules A at surface density $\rho_A^\sigma$ is

$$\begin{aligned} W_{pl/pl} &= -\frac{\pi\rho_A^\sigma\rho_B^\sigma C_{AB}}{2} \int_0^\infty \frac{2\pi r dr}{\left(D^2+r^2\right)^2} \\ &= -\frac{\pi^2\rho_A^\sigma\rho_B^\sigma C_{AB}}{2} \int_0^\infty \frac{dr^2}{\left(D^2+r^2\right)^2} = -\frac{\pi^2\rho_A^\sigma\rho_B^\sigma C_{AB}}{2D^2} \end{aligned}$$

The force per unit area is

$$f_{pl/pl} = -\frac{\pi^2\rho_A^\sigma\rho_B^\sigma C_{AB}}{D^3}$$

2. **Hamaker constant between metals**. Inserting the dispersion relation (6.26) for $\varepsilon_1$ and $\varepsilon_2$ and $\varepsilon_3 = 1$ into the integral of Eq. (6.23) leads to the integral

$$\int_{\nu_1}^\infty \left(\frac{\nu_e^2/\nu^2}{2+\nu_e^2/\nu^2}\right)^2 d\nu = \frac{\nu_e^4}{4}\int_{\nu_1}^\infty \left(\frac{1}{\nu^2+\nu_e^2/2}\right)^2 d\nu$$

$$= \frac{\nu_e^4}{4}\left[\frac{\nu}{\nu_e^2\left(\nu^2+\nu_e^2/2\right)} + \frac{\sqrt{2}}{\nu_e^3}\cdot\arctan\left(\frac{\nu\sqrt{2}}{\nu_e}\right)\right]_{\nu_1}^\infty$$

$$\approx \frac{\nu_e^4}{4}\left[\frac{\sqrt{2}\pi}{2\nu_e^3} - \frac{\nu_1}{\nu_e^2\left(\nu_1^2+\nu_e^2/2\right)} - \frac{\sqrt{2}}{\nu_e^3}\left(\frac{\nu_1\sqrt{2}}{\nu_e}\right)\right]$$

Here, we wrote $\arctan x = x - x^3/3 \pm ...$ and because $\nu_1 \ll \nu_e$ used only the linear term. This expression can be further simplified:

$$\approx \frac{\nu_e^4}{4}\left[\frac{\pi}{\sqrt{2}\nu_e^3} - \frac{2\nu_1}{\nu_e^4} - \frac{2\nu_1}{\nu_e^4}\right] \approx \frac{\pi\nu_e}{4\sqrt{2}}$$

Multiplication by $3h/4\pi$ leads to Eq. (6.27). The first term in Eq. (6.23) can be neglected because $3/4\cdot k_BT \ll (3/16\sqrt{2})\,h\nu_e$.

3. **Van der Waals energy across a polymer film**. For identical materials 1 and 2 we can express the Hamaker constant by

$$A_H \approx \frac{3}{4} k_B T \cdot \left(\frac{\varepsilon_1 - \varepsilon_3}{\varepsilon_1 + \varepsilon_3}\right)^2 + \frac{3h\nu_e}{16\sqrt{2}} \cdot \frac{\left(n_1^2 - n_3^2\right)^2}{\left(n_1^2 + n_3^2\right) \cdot \sqrt{n_1^2 + n_3^2}}$$

Inserting appropriate values (Table 6.2) and with $\nu_e = 2.7 \times 10^{15}$ Hz we get

$$\begin{aligned} A_H &\approx \frac{3}{4} \cdot 5.28 \times 10^{-21}\,\mathrm{J} \cdot \left(\frac{5.4 - 2.55}{5.4 + 2.55}\right)^2 \\ &+ \frac{3 \cdot 6.63 \times 10^{-34}\,\mathrm{Js} \cdot 2.7 \times 10^{15}}{16\sqrt{2}\,\mathrm{s}} \cdot \frac{(2.496 - 2.528)^2}{(2.50 + 2.53) \cdot \sqrt{2.50 + 2.53}} \\ &= 3.97 \times 10^{-21}\,\mathrm{J} \cdot 0.129 + 2.37 \times 10^{-19}\,\mathrm{J} \cdot 0.000045 = 0.523 \times 10^{-21}\,\mathrm{J} \end{aligned}$$

$$\begin{aligned} wA &= \frac{A_H A}{12\pi d^2} = k_B T \Rightarrow A = \frac{12\pi d^2 k_B T}{A_H} \\ A &= \frac{12\pi \cdot 1\,\mathrm{nm}^2 \cdot 5.28 \times 10^{-21}\,\mathrm{J}}{0.53 \times 10^{-21}\,\mathrm{J}} = \frac{12\pi \cdot 1\,\mathrm{nm}^2 \cdot 5.28}{0.53} = 376\,\mathrm{nm}^2 \end{aligned}$$

For silicon oxide we get

$$A_H \approx 0.97 \times 10^{-21}\,\mathrm{J} \Rightarrow A = 205\,\mathrm{nm}^2$$

4. **Polystyrene films on silicon oxide**. Inserting the appropriate values ($n_2 = \varepsilon_2 = 1$) into Eq. (6.25):

$$\begin{aligned} A_H &\approx \frac{3}{4} \cdot 4.12 \times 10^{-21}\,\mathrm{J} \cdot \left(\frac{4.7 - 2.55}{4.7 + 2.55}\right) \cdot \left(\frac{1 - 2.55}{1 + 2.55}\right) \\ &+ \frac{4.83 \times 10^{-19}\,\mathrm{J} \cdot (2.372 - 2.528) \cdot (1 - 2.528)}{\sqrt{2.372 + 2.528} \cdot \sqrt{1 + 2.528} \cdot \left(\sqrt{2.372 + 2.528} + \sqrt{1 + 2.528}\right)} \\ &= -4.0 \times 10^{-22}\,\mathrm{J} + 7.4 \times 10^{-21}\,\mathrm{J} = 7.0 \times 10^{-21}\,\mathrm{J} \end{aligned}$$

A positive Hamaker constant corresponds to an attractive force between the silicon oxide–polystyrene and the polystyrene–air interfaces. This implies that the film is not stable. If it is thin enough and has a chance, for instance when annealing, the film ruptures and holes are formed.

5. **Van der Waals force on atomic force microscope tip**. For a parabolically shaped tip we have $r^2/(2R) = x - d$. With a cross-sectional area at height $x$ of $A = \pi r^2 = 2\pi R\,(x - d)$ for $x \geq d$ we get $dA/dx = 2\pi R$ and

$$F(d) = \int_d^\infty f(x)\,\frac{dA}{dx} dx = -\int_d^\infty \frac{2\pi R A_H}{6\pi x^3} dx = -\frac{A_H R}{3} \int_d^\infty \frac{dx}{x^3} = -\frac{A_H R}{6d^2}$$

6. **Exponential force law**. Parabolic shape:

$$\begin{aligned} F(d) &= \int_d^\infty f(x)\,\frac{dA}{dx}dx = \int_d^\infty f_0 e^{-\kappa x}\cdot 2\pi R\cdot dx \\ &= -\frac{2\pi R f_0}{\kappa}\int_d^\infty e^{-\kappa x}dx = -\frac{2\pi R f_0}{\kappa}\,e^{-\kappa d} \end{aligned}$$

Conical shape:

$$\begin{aligned} F(d) &= \int_d^\infty F^A(x)\,\frac{dA}{dx}dx = \int_d^\infty f_0 e^{-\kappa x}\cdot 2\pi\cdot\tan^2\alpha\cdot(x-d)\cdot dx \\ &= 2\pi f_0\tan^2\alpha\cdot\left[\int_d^\infty x e^{-\kappa x}dx - d\int_d^\infty e^{-\kappa x}dx\right] \\ &= 2\pi f_0\tan^2\alpha\cdot\left[-\frac{e^{-\kappa x}}{\kappa^2}(\kappa x+1)+\frac{d}{\kappa}e^{-\kappa x}\right]_d^\infty = \frac{2\pi f_0\tan^2\alpha}{\kappa^2}\cdot e^{-\kappa d} \end{aligned}$$

7. **Contact radius between silica spheres**. We use $E = 5.4\times 10^{10}\,\mathrm{Pa}$ and $\nu = 0.17$. $R^* = R/2 = 10^{-6}\,\mathrm{m}$. The contact area is $\pi a^2$. The reduced Young's modulus is obtained by

$$\frac{1}{E^*} = 2\frac{1-\nu^2}{E} = \frac{2\cdot 0.971}{5.4\times 10^{10}\,\mathrm{Pa}} = 3.60\times 10^{-11}\,\mathrm{Pa}^{-1} \Rightarrow E^* = 2.8\times 10^{10}\,\mathrm{Pa}$$

Hertz model:

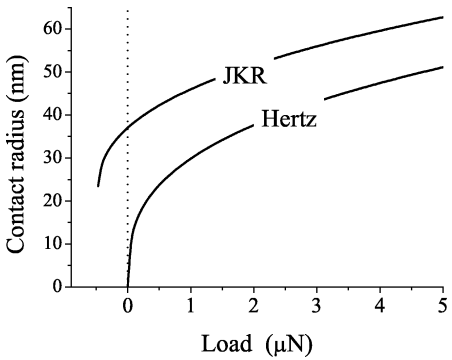
$$\begin{aligned} a &= \sqrt[3]{\frac{3R^*}{4E^*}\cdot F} = \sqrt[3]{\frac{3\times 10^{-6}\,\mathrm{m}^3}{1.12\times 10^{11}\,\mathrm{N}}\cdot F} \\ &= \sqrt[3]{2.67\times 10^{-17}\,\frac{\mathrm{m}^3}{\mathrm{N}}\cdot F} = \sqrt[3]{2.67\times 10^{4}\,\frac{\mathrm{nm}^3}{\mu\mathrm{N}}\cdot F} \end{aligned}$$

JKR model: In contrast to the Hertz model where the minimal load is zero, here we can apply negative loads, that is we can even pull on the particles. The greatest negative load is equal to the adhesion force $3\pi\gamma_S R^* = 0.471\,\mu\mathrm{N}$.

$$a^3 = \frac{3}{4}\,\frac{R^*}{E^*}\left(F + 3\pi W R^* + \sqrt{6\pi W R^* F + (3\pi W R^*)^2}\right)$$

8. **Dispersion in aqueous electrolyte**. According to Derjaguin's approximation (Eq. 6.36) the force between two particles is

$$F = \pi R w(D) = 64\pi R c_0 k_B T\lambda_D\cdot\left(\frac{e^{e\Psi_0/2k_BT}-1}{e^{e\Psi_0/2k_BT}+1}\right)^2\cdot e^{-\frac{D}{\lambda_D}} - \frac{A_H R}{12D^2}$$




The energy is

$$\begin{aligned} W(D) &= \int_D^\infty F \cdot dD = 64\pi R c_0 k_B T \lambda_D \left( \frac{e^{\frac{e\Psi_0}{2k_BT}} - 1}{e^{\frac{e\Psi_0}{2k_BT}} + 1} \right)^2 \cdot \int_D^\infty e^{-\frac{d'}{\lambda_D}} dD \\ &- \frac{A_H R}{12} \int_D^\infty \frac{1}{d'^2} dD \\ &= 64\pi R c_0 k_B T \lambda_D^2 \left( \frac{e^{\frac{e\Psi_0}{2k_BT}} - 1}{e^{\frac{e\Psi_0}{2k_BT}} + 1} \right)^2 \cdot e^{-\frac{D}{\lambda_D}} - \frac{A_H R}{12D} \end{aligned}$$

When plotting the energy versus distance for different salt concentrations we find that, at 0.28 M, we have an energy barrier of roughly $10k_BT$.

## Chapter 7: Contact angle phenomena and wetting

1. **Capillary rise**. We use Eq. (7.13) and Youngs equation:

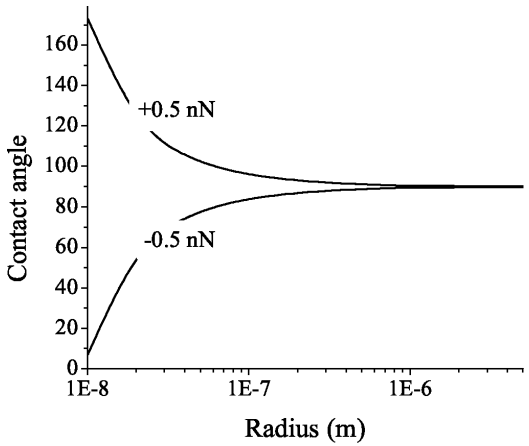
$$\gamma_S - \gamma_{SL} = \gamma_L \cos\Theta = \frac{h r_C g \rho}{2} = 0.054\ \frac{\mathrm{N}}{\mathrm{m}}$$

2. **Line tension effects**. Rearranging Eq. (7.10):

$$\cos\Theta = \frac{\gamma_S - \gamma_{SL} - \kappa/a}{\gamma_L}$$

To find $\gamma_S - \gamma_{SL}$ we use $\gamma_L \cos\Theta + \kappa/a = \gamma_S - \gamma_{SL}$. Thus at $a = 0.5\,\mathrm{mm}$ we have

$$0.05\,\frac{\mathrm{N}}{\mathrm{m}} \cdot \cos 90^\circ + \frac{0.5 \times 10^{-9}\,\mathrm{N}}{5 \times 10^{-4}\,\mathrm{m}} = 10^{-6}\,\frac{\mathrm{N}}{\mathrm{m}} = \gamma_S - \gamma_{SL}$$




3. **Surface energy of polymers**

$$\gamma_{SL} = \gamma_S + \gamma_L - 2\sqrt{\gamma_S^d \gamma_L^d} - 2\sqrt{\gamma_S^p \gamma_L^p} \Rightarrow \gamma_L = \gamma_{SL} - \gamma_S + 2\sqrt{\gamma_S^d \gamma_L^d} + 2\sqrt{\gamma_S^p \gamma_L^p}$$

Inserting Young's equation $\gamma_L \cos\Theta = \gamma_S - \gamma_{SL}$ leads to

$$\gamma_L = -\gamma_L \cos\Theta + 2\sqrt{\gamma_S^d \gamma_L^d} + 2\sqrt{\gamma_S^p \gamma_L^p} \Rightarrow \frac{\gamma_L}{2}(\cos\Theta + 1) = \sqrt{\gamma_S^d \gamma_L^d} + \sqrt{\gamma_S^p \gamma_L^p}$$

We can solve this if we have at least two sets of results (with different liquids). For polyethylene we use (without units):

Water: $\frac{72.8}{2}(1 + \cos 94°) = 33.86 = \sqrt{\gamma_S^d 21.8} + \sqrt{\gamma_S^p 51.0}$

MI: $\frac{50.8}{2}(1 + \cos 52°) = 41.04 = \sqrt{\gamma_S^d 49.5} + \sqrt{\gamma_S^p 1.3}$

Multiplying the last equation by $\sqrt{21.8/49.5} = 0.6636$ and subtracting the second from the first equation, leads to

$$33.86 - 41.04 \cdot 0.6636 = 6.63 = \sqrt{\gamma_S^p 51.0} - 0.6636\sqrt{\gamma_S^p 1.3}$$

$$= 6.38 \cdot \sqrt{\gamma_S^p} \Rightarrow \gamma_S^p = \left(\frac{6.63}{6.38}\right)^2 = 1.08\frac{\mathrm{mN}}{\mathrm{m}}$$

$$33.86 = \sqrt{\gamma_S^d 21.8} + \sqrt{1.08 \cdot 51.0} \Rightarrow$$

$$\gamma_S^d = \frac{\left(33.86 - \sqrt{1.08 \cdot 51.0}\right)^2}{21.8} = 32.06\frac{\mathrm{mN}}{\mathrm{m}}$$

$$\gamma_S = \gamma_S^d + \gamma_S^p = (32.06 + 1.08)\frac{\mathrm{mN}}{\mathrm{m}} = 33.1\frac{\mathrm{mN}}{\mathrm{m}}$$

For poly(vinyl chloride):

Water $\frac{72.8}{2}\,(1+\cos 87^\circ) = 38.3 = \sqrt{\gamma_S^d 21.8} + \sqrt{\gamma_S^p 51.0}$

MI $\frac{50.8}{2}\,(1+\cos 36^\circ) = 45.9 = \sqrt{\gamma_S^d 49.5} + \sqrt{\gamma_S^p 1.3}$

Multiplying the last equation by $\sqrt{21.8/49.5} = 0.6636$ and subtracting the second from the first equation leads to

$$38.3 - 45.9 \cdot 0.6636 = 7.84 = 6.38 \cdot \sqrt{\gamma_S^p} \Rightarrow$$

$$\gamma_S^p = \left(\frac{7.84}{6.38}\right)^2 = 1.51\frac{\mathrm{mN}}{\mathrm{m}}$$

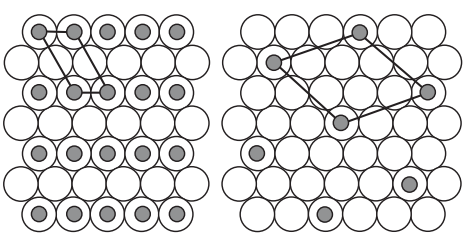
$$38.3 = \sqrt{\gamma_S^d 21.8} + \sqrt{1.51 \cdot 51.0} \Rightarrow \gamma_S^d = \frac{\left(38.3 - \sqrt{1.51 \cdot 51.0}\right)^2}{21.8} = 40.0\frac{\mathrm{mN}}{\mathrm{m}}$$

$$\gamma_S = \gamma_S^d + \gamma_S^p = (40.0 + 1.5)\frac{\mathrm{mN}}{\mathrm{m}} = 41.5\frac{\mathrm{mN}}{\mathrm{m}}$$

In a similar way we get $\gamma_S = 42.1$ mN/m for polystyrene and $\gamma_S = 40.0$ mN/m for PMMA.

## Chapter 8: Solid surfaces

1. **Overlayer structures on a fcc(111) surface**.



(1×2) (√7×√7) R 19.1°

2. **Force in atomic force microscopy**. It is lower because the van der Waals attraction is reduced and the meniscus force is absent.

3. **Structure factor of diamond**. The diamond lattice can be viewed as a fcc lattice with a basis of $c_1 = 0$, $c_2 = (a/4)(\vec{e}_x + \vec{e}_y + \vec{e}_z)$ where $a$ is the lattice constant of the cubic lattice and $\vec{e}_x$, $\vec{e}_y$, $\vec{e}_z$ are unit vectors along the cubic axes. The primitive vectors of the fcc lattice are (Fig. 14.1)

$$\vec{a_1} = \frac{a}{2}(\vec{e}_y + \vec{e}_z), \quad \vec{a_2} = \frac{a}{2}(\vec{e}_x + \vec{e}_z), \quad \vec{a_3} = \frac{a}{2}(\vec{e}_x + \vec{e}_y)$$

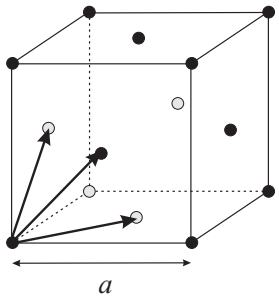



**Figure 14.1:** Primitive vectors of the fcc lattice.

The lattice vectors of the reciprocal lattice are given by

$$\vec{b}_1 = 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3} = \frac{2\pi}{a}(\vec{e}_y + \vec{e}_z - \vec{e}_x)$$
$$\vec{b}_2 = 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3} = \frac{2\pi}{a}(\vec{e}_x + \vec{e}_z - \vec{e}_y)$$
$$\vec{b}_3 = 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3} = \frac{2\pi}{a}(\vec{e}_x + \vec{e}_y - \vec{e}_z)$$

Then we get:

$$S_G = 1 + \exp(-i\vec{q} \cdot \vec{d}_2) \qquad \text{with} \qquad \vec{q} = n_1\vec{b}_1 + n_2\vec{b}_2 + n_3\vec{b}_3$$

$$\begin{aligned} \vec{q} \cdot \vec{d}_2 &= \frac{a}{4}(n_1\vec{b}_1 + n_2\vec{b}_2 + n_3\vec{b}_3) \cdot (\vec{e}_x + \vec{e}_y + \vec{e}_z) \\ &= \frac{\pi}{2}(n_1 + n_1 - n_1 + n_2 - n_2 + n_2 + n_3 + n_3 - n_3) \\ &= \frac{\pi}{2}(n_1 + n_2 + n_3) \end{aligned}$$

$$S_G = 1 + \exp\left[-i\frac{\pi}{2}(n_1 + n_2 + n_3)\right] = 1 + (-1)^{\frac{1}{2}(n_1+n_2+n_3)}$$

| | | | |
|---|---|---|---|
| $S_G = 2 \pm i$ | for | $n_1 + n_2 + n_3$ | twice an even number |
| $S_G = 1 \pm i$ | for | $n_1 + n_2 + n_3$ | odd |
| $S_G = 0 \pm i$ | for | $n_1 + n_2 + n_3$ | twice an odd number |

4. **Grazing incidence X-ray diffraction**. With an energy $E = h\nu = hc/\lambda$ we get a wavelength

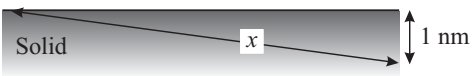
$$\lambda = \frac{hc}{E} = \frac{6.63 \times 10^{-34}\,\mathrm{Js} \cdot 3.00 \times 10^{8}\,\mathrm{m/s}}{10000 \cdot 1.60 \times 10^{-19}\,\mathrm{J}} = 1.24\,\text{Å}$$

This is just sufficiently small to analyze atomic structures.

$$\frac{\mu}{\rho} = 691\frac{\mathrm{cm}^2}{\mathrm{g}} \Rightarrow \mu = 691\frac{\mathrm{cm}^2}{\mathrm{g}} \cdot 12.0\frac{\mathrm{g}}{\mathrm{cm}^3} = 8292\,\mathrm{cm}^{-1}$$

$$e^{-\mu x} = 0.8 \Rightarrow x = -\frac{\ln 0.8}{\mu} = \frac{0.223}{8292}\,\text{cm} = 0.269\,\mu\text{m}$$

$$\tan\vartheta = \frac{10^{-9}\,\text{m}}{0.269 \times 10^{-6}\,\text{m}} \Rightarrow \vartheta = 0.21^\circ$$




## Chapter 9: Adsorption

1. **Significance of Langmuir parameter**. Linear increase of $\theta$-versus-$P$ for low pressures. For low $P$ we have $K_L P \ll 1$ so that

$$\theta = \frac{K_L P}{1 + K_L P} \approx K_L P$$

Thus, for $P \to 0$ we get $d\theta/dP = K_L$.

Half coverage is reached at a pressure of

$$\frac{1}{2} = \frac{K_L P_{1/2}}{1 + K_L P_{1/2}} \Rightarrow \frac{1}{2} + \frac{K_L P_{1/2}}{2} = K_L P_{1/2} \Rightarrow \frac{1}{2} = \frac{K_L P_{1/2}}{2} \Rightarrow P_{1/2} = \frac{1}{K_L}$$

2. **Langmuir parameter**.

$$\begin{aligned} K_L^0 &= \frac{6.02 \times 10^{23}\,\text{mol}^{-1} \cdot 16 \times 10^{-20}\,\text{m}^2 \cdot 10^{-12}\,\text{s}}{\sqrt{2\pi \cdot 0.028\,\text{kg mol}^{-1} \cdot 8.31\,\text{JK}^{-1}\,\text{mol}^{-1} \cdot 79\,\text{K}}} \\ &= 8.96 \times 10^{-9}\,\frac{\text{ms}^2}{\text{kg}} = 8.96 \times 10^{-9}\,\text{Pa}^{-1} \end{aligned}$$

$$K_L = K_L^0 \cdot e^{Q/RT} = 8.96 \times 10^{-9}\,\text{Pa}^{-1} \cdot e^{\frac{2000 \cdot 4.184}{8.31 \cdot 79}} = 3.10 \times 10^{-3}\,\text{Pa}^{-1}$$

3. **BET adsorption isotherm equation**. From Eq. 9.36 we get

$$\frac{k_{ad}^1}{a_1} = \frac{S_1}{S_0 P} e^{-Q_1/RT}$$

In thermodynamic equilibrium this is equal to $1/P$ because the Boltzmann factor $e^{-Q_1/RT}$ determines the occupancy of the two states and we have $S_1/S_0 = e^{Q_1/RT}$. With the same argument we get $k_{ad}^i/a_i \approx 1/P$.

4. **Hydraulic number for close-packed spheres**. We consider a volume $V$ which contains $n$ spheres. Here, $n$ is assumed to be a large number so that it is practically continuous. The volume filled by particles is $V_f = n \cdot 4/3 \cdot \pi R^3$. The total surface area of all particles is $A = n \cdot 4\pi R^2$. The free ("void") volume is $V_v = 0.26\,V = 0.26 \cdot n \cdot 4/3 \cdot \pi R^3/0.74$. Thus the hydraulic number is

$$a_h = \frac{V_v}{A} = \frac{0.26 \cdot n \cdot 4/3 \cdot \pi R^3}{0.74 \cdot n \cdot 4\pi R^2} = \frac{0.26}{0.74 \cdot 3} R = 0.117 \cdot R$$

5. **Docosane on graphite**. The number of molecules per unit area is

$$\frac{88.9 \times 10^{-6}\,\mathrm{mol}}{68\,\mathrm{m}^2} = 1.31 \times 10^{-6} \frac{\mathrm{mol}}{\mathrm{m}^2} = 7.89 \times 10^{17}\,\mathrm{m}^{-2}$$

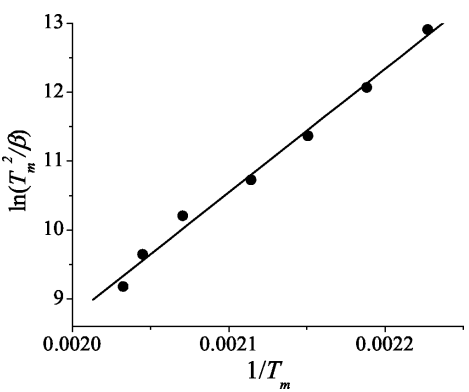
The area per molecule is $1/7.89 \times 10^{17}\,\mathrm{m}^{-2} = 1.27 \times 10^{-18}\,\mathrm{m}^2 = 1.27\,\mathrm{nm}^2$. The cross-sectional area of an alkane chain is roughly 4.5 Å × 4.5 Å = 0.2 nm$^2$. The area occupied by a molecule lying flat on the surface is roughly 4.5 Å × 1.2 Å × 22 = 1.2 nm$^2$. Conclusion: The molecules are lying flat on the surface.

6. **Thermal desorption spectroscopy**.

$$a \cdot e^{-E_{des}/RT_m} = \frac{E_{des}\beta}{RT_m^2} \Rightarrow \frac{\beta}{T_m^2} = \frac{aR}{E_{des}} \cdot e^{-E_{des}/RT_m} \Rightarrow$$

$$\ln \frac{\beta}{T_m^2} = \ln \frac{aR}{E_{des}} - \frac{E_{des}}{RT_m} \ln\left(\frac{T_m^2}{\beta}\right) = \frac{E_{des}}{RT_m} + \ln\left(\frac{E_{des}}{aR}\right)$$

If we now plot $\ln\left(T_m^2/\beta\right)$ versus $1/T_m$ we should get a straight line with a slope $E_{des}/R$ and an intercept at $\ln\left(E_{des}/(aR)\right)$.




From a linear fit of the graph we get

$$\frac{E_{des}}{R} = 17900 \Rightarrow E_{des} = 149\,\frac{kJ}{mol}$$

$$\ln\left(\frac{E_{des}}{aR}\right) = -27.1 \Rightarrow a = \frac{E_{des}}{R} \cdot e^{27.1} = 1.1 \times 10^{16}\,\mathrm{Hz}$$

## Chapter 10: Surface modification

1. **Hydrophobization/hydrophilization of silicon**. a) To hydrophobize silicon surfaces you first have to hydrophilize the silicon to offer a high hydroxy group density for the consecutive coupling reaction of the silanes. Then it is cleaned and water is removed by washing with different organic solvents with decreasing polarity. Afterwards the silane, e.g. octadecyltrichlorsilane (OTS) is dissolved in a water-free organic solvent. The silicon substrate is immersed for some hours and finally washed again with organic solvents with increasing polarity to get rid of physisorbed silanes.

   b) Form a thiol monolayer with a thiol having a hydrophilic rest group such as 11-mercaptoundecanol or 11-mercaptoundecanoic acid. Both can be dissolved in ethanol.

   c) A first cross-check for the proper surface modification is to measure the contact angle of a water drop on top of the substrate. If the contact angle is above 90° the surface can be called hydrophobic. Hydrophilic surface should have contact angles close to zero. For comparison, the contact angle on the substrate before surface modification should also be measured. A test for surface homogeneity is the contact angle hysteresis. A perfectly smooth, homogeneous surface should show low hysteresis (see Section 7.3.2). Quantitative measures of film thickness are provided by ellipsometry. Monolayer thickness can be compared to the expected molecular length of the adsorbates. More details can be measured by grazing incidence X-ray diffraction. They provide film thickness, electron density perpendicular to the surface and molecular orientation within the adsorbate layer. Atomic force microscopy or scanning electron microscopy provide information about the homogeneity or roughness of the surface layer up to a molecular resolution. Also, mechanical properties like the elasticity can be evolved from surface force measurements with the AFM. The denser the adsorbate layer, the stiffer it appears.

2. a) One driving force for protein folding and structure is the hydrophobic effect. Almost all proteins are in an aqueous environment. The folding brings most hydrophobic groups into the interior of the protein while hydrophilic groups are on the surface. The three-dimensional structure of proteins is further maintained by hydrogen-bonds between the peptide groups. Additionally, the unspecific van der Waals interaction between different parts of the amino acids, leads to an attraction. Moreover, there might also be an attractive or repelling interaction between charges in the amino acids.

   When a protein adsorbes on a surface, the above interactions will be perturbed. For example, if the surface is hydrophobic the protein usually unfolds to bring its hydrophobic grous into contact with the surface. This often leads to a denaturing of the protein; it will be changed irreversibly and no longer work properly.

   b) In order to prevent adsorption, the solid surface might be covered with hydrophilic polymer brushes. They must be so long that they tend to assume a coil structure, but they must be so short that they still can be ordered to brushes. In this case there will be an entropic repulsion when the protein approaches the surface. It tries to order the polymer brushes on top of the surface which tries to escape from this unfavorable situation. For details see Ref. [449].

3. a) Glucose is oxidated to gluconic acid ($CH_2OH$-$(CHOH)_4$-COOH). An hydroxy group is built in and the hydrogen atom is taken away, together with two electrons:

$$\text{Glucose} + H_2O \rightarrow \text{Gluconic acid} + 2H^+ + 2e^-$$

We take the hydroxy group from the water molecule and a positively charged hydrogen ion, a proton $H^+$ is left over. Together, we have left over, two protons and two electrons on the right side of the equation. For reduction there are two hydrogen atoms too few on the left side:

$$O_2 + 2H^+ + 2e^- \rightarrow H_2O_2$$

In water there are no single hydrogen atoms or hydrogen molecules, $H_2$. Therefore, we have to take two protons on the left side to have a correct mass balance. Now we have two positive charges on the left side but no charges on the right side. Therefore, we have to add two electrons to the left side and the mass and charge balance is satisfied. Putting together the two single equations to one net equation we state that we can omit the protons as well as the electrons because they occur on both sides of the equation:

$$\text{Glucose} + H_2O + O_2 \rightarrow \text{Gluconic acid} + H_2O_2$$

This procedure for chemical equations is very similar to mathematical or physical equations. Please note that the actual reaction is much more complicated. For example, the hydroxy group in the gluconic acid does not come from the water molecule. Glucose in fact usually forms a ring structure instead of the shown linear structure.

b) Depending on your method of measurement you might have the magnitudes of different physical properties to measure. Gluconic acid is a charged molecule, so the *conductivity* changes during the reaction when carried out in solution. Hydrogen peroxide forms with progress of the reaction. It is a reactive chemical. You might know it from hair dyeing . So, it can easily be reduced to water, oxidizing another sensor molecule to a colored form. This could be monitored optically by *light adsorption*. Another possibility is the oxidation of hydrogen peroxide back to oxygen. Something has to take over the emitting electrons. This can be an electrode. So if you immobilise glucose oxidase close to an electrode without denaturing the protein, the electrode will quantitatively measure the content of sugar by means of electrochemistry. An overview of biosensors and their economic impact is Ref. [406].

4. We must have single electrodes in the system in order to derive an electrical signal, so let us take gold. It can be deposited on the polystyrene through masks by sputtering. Neurons like soft, hydrophilic surfaces such as polyelectrolytes which also contain a lot of water. How do we adsorb polyelectrolytes on the more hydrophobic gold (see Table 7.1) without the risk of washing the polymers away during cultivation? We have to hydrophilize the gold, e.g., by thiols containing hydrophilic end-groups. Usually the end groups must be capable of covalently binding the polyelectrolyte on top.

5. The complete design scheme is shown in the following sketch. The first step is an oxidation of the silicon surface at an elevated temperature (some 100°C) to get a thickness of

some 100 nm (thermal oxidation). The native silicon oxide layer is only about 1 to 2 nm thick. This oxidation layer later serves as an insulator. The second step is sputtering of a chromium layer on the oxidized silicon. Then gold is sputtered on top. The chromium layer increases the adhesion of gold to the silicon substrate. Now, we have to think of how to get the gold and chromium layer interrupted. Metals, especially gold, can be removed by chemicals. So we apply the isotropic etching. For gold it is carried out with potassium iodine/iodine 1:1 and not with aquia regia for reasons of selectivity. But how to get the local dissolution? A mask must be applied. You cannot use just a mobile mask because the etching solution would creep between mask and device. Therefore, we have to deposit a more resistant layer with strong adhesion to the gold. On the other hand, the mask must be completely removed after the gold etching step. So, organic material like polymers are perfect for this purpose.

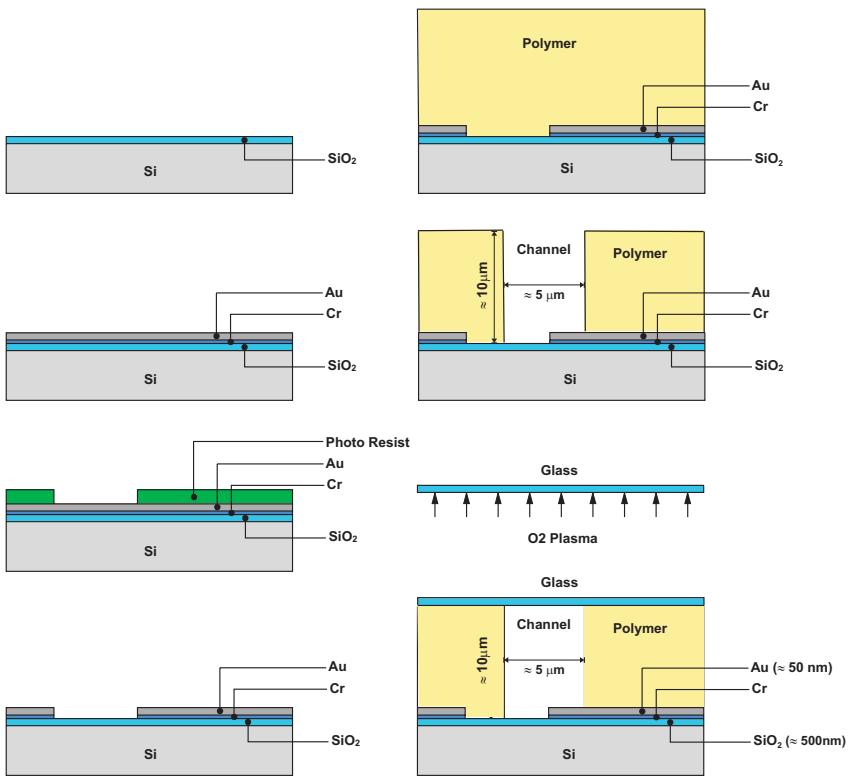



**Figure 14.2:** The figure was taken by permission of M. Böhm.

A photoresist is used which is sensitive to radiation. The radiation polymerizes the spin-coated material locally (light can easily be focussed on a desired spot by masks) and stabilizes it against the following etching step. After radiation a washing step is included to remove the unpolymerized part (the hole in this case). Afterwards the gold and chromium layers are chemically etched and the polymer mask is removed completely, chemically.

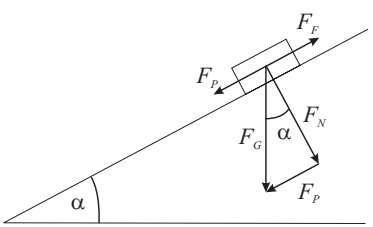
In fact, the chemical etching is isotropic which is unfavorable for the structure we want to obtain. But the gold layer thickness is in the order of some 10 nm which is too thin for the isotropy to occur.

In the next step we are form the walls of the channels in the microfluidic device. A new, very special polymer is spin-coated on the substrate to the desired thickness. This polymer differs from the inexpensive photoresist because it comes into contact with the later fluid. Therefore, it should have a long stability; it should not form cracks, should be stable against different chemicals and it should be hydrophilic or easily hydrophilized, because otherwise water will not run through the channel. Again a photo-sensitive material is used, but this time the later channel is photochemically modified. A perfect material to use is SU-8. For details see Refs. [450, 451]. This part can be washed away afterwards.

In the last step we have to put the glass cover on top. This is quite tricky because it must close any gap to the underlying polymer on an Angstrom level. This is achieved by hydrophilizing the glass substrate on one side, e.g., with oxygen plasma. After contacting the polymer surface the glass substrate cannot be moved any more because of the work of adhesion; the back formation of the bare glass surface costs too much energy. Additionally, the glass surface inside the channel is hydrophilic now, thus supporting the transport of water.

## Chapter 11: Friction, lubrication, and wear

1. **Sliding on an inclined plane**. The gravitational force $F_G$ acting on the block can




be deconstructed into a force $F_N$ normal to the surface and a force $F_P$ acting parallel to the surface. Looking at the absolute values of the forces, trigonometry tells us that $F_P = \tan\alpha F_N$. From Amontons' law we get $F_R = \mu_s F_N$. When the body starts sliding, the absolute value of the force parallel to the plane must equal the friction force. From this we get:

$$\mu_s = \tan\alpha \Rightarrow \alpha = \arctan\mu = \arctan 0.5 = 26.6^\circ$$

When the bock is sliding, $F_P$ stays the same, but the friction force is reduced to $F_R = \mu_k F_N$. The net force acting parallel to the surface is therefore:

$$F'_P = F_P - F'_R = \mu_s F_N - \mu_k F_N = 0.1 F_N = 0.1 \sin\alpha \cdot F_G = 0.1 mg \sin\alpha$$

The acceleration is given by $a = \frac{F}{m}$ and the speed after a distance $L$ is equal to

$$\begin{aligned} v &= \sqrt{2aL} = \sqrt{2\frac{F}{m}L} = \sqrt{2\frac{0.1\sin\alpha \cdot mg}{m}L} = \sqrt{2 \cdot 0.1\cos\alpha \cdot gL} \\ &= \sqrt{0.2\cos 26.6 \cdot 9.81\frac{\mathrm{m}}{\mathrm{s}^2} \cdot 1\,\mathrm{m}} = 1.32\,\frac{\mathrm{m}}{\mathrm{s}} \end{aligned}$$

2. **True contact area**. The contact surface will deform, until the pressure is equal to the yield stress $y$. If $G$ is the weight force and $A$ the true contact area, $\rho$ the density and V the volume of the cube, we can write:

$$\begin{aligned} y &= \frac{G}{A} = \frac{\rho V g}{A} \\ A &= \frac{\rho V g}{y} = \frac{7.85\,\mathrm{kg/cm^3} \cdot (10\,\mathrm{cm})^3 \cdot 9.81\,\mathrm{m/s^2}}{10^9\,\mathrm{N/m^2}} \\ &= 7.7 \cdot 10^{-8}\,\mathrm{m}^2 = 7.7 \cdot 10^4\,\mu\mathrm{m}^2 \end{aligned}$$

This corresponds to 7700 microcontacts of $10\,\mu\mathrm{m}^2$ each.

3. **Hydrodynamic lubrication**. The sliding speed is $v = 2\pi \cdot 0.005\,\mathrm{m} \cdot 80\,\mathrm{Hz} = 1.26\,\mathrm{m/s}$. The friction force is $F_F = \frac{0.1\,\mathrm{m} \cdot 2\pi \cdot 0.005\,\mathrm{m} \cdot 1.26\,\mathrm{m/s} \cdot 2.49 \times 10^{-3}\,\mathrm{Pas}}{3 \times 10^{-6}\,\mathrm{m}} = 3.29\,\mathrm{N}$. The torque is $M = 3.29\,\mathrm{N} \cdot 0.005\,\mathrm{m} = 0.0164\,\mathrm{Nm}$.

## Chapter 12: Surfactants, micelles, emulsions, and foams

1. **Hydrophobic effect drives micellization**. We take two alkylethylene glycols with constant head group but different chain length, namely $C_8E_6$ and $C_{12}E_6$. Their CMCs are 9.8 mM and 0.08 mM, respectively. They are nonionic so Eq. (12.3) should be valid. The Gibbs energies of micellization are $\Delta G_m^{mic} = RT \ln 0.0098 = -11.5$ kJ/mol and –23.4 kJ/mol, respectively. Division by 4 leads to a value of 3.0 kJ/mol. In total the Gibbs free energy for a transfer of an alkyl chain into the interior of a micelle is approximately $\Delta G_m^{mic}(HP) = -(n_C - 1) \cdot 3.0 - 9.6$ kJ/mol.

2. **Volume of hydrocarbon chain**. We assume that the volume $V_C$ of the $n$-alkanes is the sum of the volumes of all methylene ($\sim CH_2 \sim$) units plus twice the volume of a methyl group ($\sim CH_3$):

$$\rho = \frac{M/N_A}{n_C V_{CH_2} + 2 \cdot V_{Me}} \Rightarrow n_C V_{CH_2} + 2 \cdot V_{Me} = \frac{M}{\rho N_A}$$

Using the values for hexane and dodecane we get two equations:

$$4 \cdot V_{CH_2} + 2 \cdot V_{Me} = \frac{86.18 \times 10^{-3}\,\mathrm{kg\,mol^{-1}}}{654.8\,\mathrm{kg\,m^{-3}} \cdot 6.022 \times 10^{23}\,\mathrm{mol^{-1}}} = 0.2185\,\mathrm{nm}^3$$

$$10 \cdot V_{CH_2} + 2 \cdot V_{Me} = \frac{170.34 \times 10^{-3}\,\mathrm{kg\,mol^{-1}}}{748.7\,\mathrm{kg\,m^{-3}} \cdot 6.022 \times 10^{23}\,\mathrm{mol^{-1}}} = 0.3778\,\mathrm{nm^3}$$

Subtracting the two equations leads to

$$6 \cdot V_{CH_2} = 0.1593\,\mathrm{nm^3} \Rightarrow V_C = 0.02655\,\mathrm{nm^3}$$

By inserting this value into one of the two equations we get $V_{Me} = 0.05615\,\mathrm{nm^3}$. Using the total number of carbon atoms $n_C$ as a parameter, we get for a single, saturated alkyl chain $V_C = (0.02655 \cdot n_C + 0.0296)\ \mathrm{nm^3}$.

3. **Gibbs free energy of emulsion formation**. The Gibbs free energy is $\Delta G_{em} = \gamma A$, where $A$ is the total surface area of all drops. This area is $A = n \cdot 4\pi R^2$, with $n$ being the number of drops. With the total volume $V = n \cdot 4\pi R^3/3 \Rightarrow n = 3V/4\pi R^3$ we get $\Delta G_{em} = 3\gamma V/R$.

4. **Two merging drops**. For constant volume we have $2\frac{4}{3}\pi r^3 = \frac{4}{3}\pi r_{new}^3 \Rightarrow r_{new} = \sqrt[3]{2} \cdot r$.

$$\frac{A_{old}}{A_{new}} = \frac{2 \cdot 4\pi r^2}{4\pi r_{new}^2} = \frac{2}{\sqrt[3]{2^2}} = \sqrt[3]{2} = 1.26$$

5. **Rate of flocculation**. The concentration is $c_0 = 10^{15}$ droplets per m$^3$. With $k_f = 6.2 \times 10^{-19}\,\mathrm{m^3\,s^{-1}}$ the initial decrease in concentration is $-\frac{dc}{dt} = k_f c^2 = 6.2 \times 10^{-19}\,\frac{\mathrm{m^3}}{\mathrm{s}}\left(10^{15}\,\mathrm{m^{-3}}\right)^2 = 6.2 \times 10^{11}$ droplets per m$^3$ per second. So, we can estimate that within the first second the relative decrease in concentration is only 0.06%.

6. **Bending energy**. For the bending energy we get

$$g^\sigma = \frac{1}{2}k_B\left(\frac{2}{R}\right)^2 = \frac{2k_B}{R^2} = \frac{20k_BT}{R^2}$$

$$\begin{aligned} R &= 5\,\mathrm{nm} \Rightarrow g^\sigma = 3.3\,\mathrm{mN/m} \\ R &= 20\,\mathrm{nm} \Rightarrow g^\sigma = 0.21\,\mathrm{mN/m} \\ R &= 100\,\mathrm{nm} \Rightarrow g^\sigma = 0.0082\,\mathrm{mN/m} \end{aligned}$$

## Chapter 13: Thin films on surfaces of liquids

1. **Pressure in monolayer**. From $P = \pi/d$ we can calculate the corresponding three-dimensional pressure as $P = \frac{0.03\mathrm{N/m}}{25 \cdot 10^{-10}\mathrm{m}} = 12 \cdot 10^6\mathrm{N/m^2} \approx 120\,\mathrm{atm}$.

2. **Transition enthalpy and entropy**. For the Clapeyron equation we applied the transformation of pressure $P$ to lateral pressure $\pi$ and volume $V$ to area $\sigma_A$. To get the molar transition enthalpies the equation has to be resolved with respect to $\Delta H_c$:

$$\Delta H_c = -\frac{\partial \pi_c}{\partial T} \cdot T\Delta\sigma_c \tag{14.1}$$

Now we have to determine the area change in the plateau of the phase transition as indicated in Fig. 13.4 for 30°C. The corresponding temperatures have to be transformed into Kelvin. To determine $\partial\pi_c/\partial T$ you can plot $\pi_c$ versus $T$, make a linear regression analysis, and take the slope. It turns out to be $\partial\pi_c/\partial T = (1.1 \pm 0.2) \times 10^{-3}\,\mathrm{Nm^{-1}K^{-1}}$. We extract the transition pressures and molecular areas from Fig. 13.4 insert the numbers into Eq. (14.1) and get the enthalpies. Division by temperature leads to the entropies.

| T (K) | $\pi_c$ (mN/m) | $\Delta\sigma_A$ (Å$^2$) | $\Delta H_c$ (kJ/mol) | $\Delta S_c$ (kJ/K/mol) |
|---|---|---|---|---|
| 293.2 | 5 | $52 - 77 = -25$ | 49 | 0.17 |
| 298.2 | 9 | $50 - 68 = -18$ | 36 | 0.12 |
| 303.2 | 16 | $50 - 64.5 = -14.5$ | 29 | 0.096 |

Two remarks: (1) The area change is negative because the area after the phase transition is lower than before. (2) The enthalpy of condensation is of the order of the standard enthalpies of vaporization of hydrocarbons.

3. **X-ray reflection**.

   (a) To calculate the thickness from the angle we use the Bragg equation (13.8). Curve a: $\alpha_1 = 1.75° \Rightarrow d_1 = 25\,\text{Å}$. Curve b: $\alpha_2 = 1.25° \Rightarrow d_2 = 35\,\text{Å}$. The thickness increases with increasing surface pressure. At higher pressure the thickness is equal to the molecular length as calculated from the structure formula ($d = 9\,\text{Å} + 10 \cdot 2.5\,\text{Å} = 34\,\text{Å}$). The molecules are oriented normal to the water surface.

   (b) The effective volume is the product of molecular area times thickness: $V = \sigma_A \cdot d$. Curve a: $V_1 = 25 \times 110\,\text{Å}^3 = 2750\,\text{Å}^3$. Curve b: $V_2 = 35 \times 35\,\text{Å}^3 = 1225\,\text{Å}^3$. The effective volume drastically decreases with pressure. Reasons: The lipid alkyl chains change from a disordered to nearly fully stretched (all-trans) conformation at high surface pressure. In the disordered phase there is still a high fraction of free volume within the monolayer. Additionally, water molecules bound to the sugar headgroup are expelled with compression (dehydration).

   (c) The theoretical volume of a water molecule can simply be calculated from the density of water ($\rho = 1.0\,\mathrm{g/cm^3}$) by

   $$V = \frac{1}{\rho_{H_2O}} \cdot \frac{M_{H_2O}}{N_A} = 30\,\text{Å}^3$$

   with $M_{H_2O}$= 18 g/mol being the molar mass of water. From the electron density profile we get an electron density of 0.33 e/Å$^3$. The reciprocal is the volume of one electron in a water molecule. Water has 10 electrons. Thus the volume of one water molecule is 30 Å$^3$, which is the correct value.

   (d) $d_1 = 14\,\text{Å}$; $d_2 = 9\,\text{Å}$; $\rho_1 = 0.29\,\mathrm{e/Å^3}$; $\rho_2 = 0.37\,\mathrm{e/Å^3}$. Film 1 can be interpreted as the apolar alkyl chain part of the monolayer. Film 2 corresponds to the layer of polar headgroup. The electron density 1 is typical for a liquid-expanded alkyl chain layer. For a liquid-condensed alkyl layer it would be close to that of water. The headgroup contains relatively heavy oxygen atoms. Therefore, its electron density is higher.

(e) From the molecular structure we count 444 electrons per molecule. From the electron density we calculate a number of

$$n_e = \left(14\,\text{Å} \cdot 0.29\,\text{e/Å}^3 + 9\,\text{Å} \cdot 0.37\,\text{e/Å}^3\right) \cdot 110\,\text{Å}^2 = 813\,\text{electrons}$$

The difference can be attributed to water molecules in the headgroup regime. The difference of about 370 electrons which represent 37 water molecules (10 electrons/water molecule) appears too high. No measurement is perfect!

4. **Optical interference**. We assume that the light is coupled perpendicular into the thin film ($\alpha = 90°$). Case 1: $1 < n_1 < n_2$. The phase shift due to a difference in optical path length between the beams reflected at the two interfaces 1 and 2 is $n_1 2d$. The reflection at the interface between air and $n_1$ leads to a phase shift of $\pi$, as well as the reflection at the interface between $n_1$ and $n_2$. Thus, it is identical for both beams. The condition for the first minimum to occur (destructive interference) is that the total phase shift between 1 and 2 should be $\lambda/2$. Thus we get:

$$2n_1 d = \lambda/2 \qquad \text{or} \qquad d = \lambda/(4n_1)$$

Case 2: $1 < n_1 > n_2$, i.e., the optical density of the thin film is higher than that of the surrounding media. Now we have to consider the phase shift at the interface between $n_1$ and $n_2$. The total phase shift is $2n_1 d - \pi$ or $2n_1 d - \lambda/2$. For the first minimum we have a phase shift of $\lambda/2$ which leads to

$$2n_1 d - \frac{\lambda}{2} = \frac{\lambda}{2} \qquad \text{or} \qquad d = \frac{\lambda}{2n_1}$$

# A Analysis of diffraction patterns

Here we introduce the more general theory of diffraction at crystal surfaces. First, we analyze for which directions of the outgoing radiation we get constructive interference and observe "diffraction peaks". In the second part we discuss what determines the intensities of these maxima.

## A.1 Diffraction at three dimensional crystals

### A.1.1 Bragg condition

A simple framework to understand the diffraction of electromagnetic radiation or particles by periodic structures has been introduced by Bragg[1]. In this model, we look for the interference of rays that are scattered from different parallel crystal planes with a spacing $d$ (Fig. A.1). We first treat the case of specular reflection. In specular reflection the angle between incident rays and the lattice plane is the same as the angle between outgoing rays and the lattice plane. The condition for constructive interference is given by the Bragg equation:

$$n\lambda = 2d \sin \vartheta, \tag{A.1}$$

where $n$ is an integer that is called the order of the diffraction peak. If we measure the angles under which diffraction peaks occur, we can determine the distances between the lattice planes. In the case of a two-dimensional array (crystal surface) or a three-dimensional periodic arrangement (bulk crystal), we have to fulfill a Bragg equation for each dimension simultaneously. Constructive interference occurs only at angles were the diffraction patterns of different dimensions overlap with each other. This strongly limits the number of observed diffraction peaks.

### A.1.2 Laue condition

To discuss diffraction for the more general case, where the condition of specular reflection is given up, it is convenient to introduce the so-called wave vector. The direction of the wave vector $\vec{k}$ is that of the propagating wave and its length is given by

$$|\vec{k}| = \frac{2\pi}{\lambda} \tag{A.2}$$

[1] William H. Bragg ,1862–1942, professor of mathematics and physics and his son W. Lawrence Bragg, 1890–1971, professor of physics, received together the Nobel Prize for Physics in 1915.

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

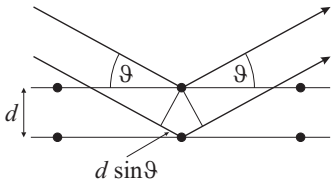



**Figure A.1:** Scattering of radiation at parallel lattice planes.

Instead of the scattering angle $\vartheta$ we use the so-called scattering vector $\vec{q}$. The scattering vector is the difference between the incoming wave vector $\vec{k}_i$ and the outgoing wave vector $\vec{k}_f$:

$$\vec{q} = \vec{k}_f - \vec{k}_i. \tag{A.3}$$

Since we assume elastic scattering (no change in energy and wavelength), the norm of the incident wave vector is equal to that of the outgoing wave vector, $|\vec{k}_i| = |\vec{k}_f|$, and the maximum length of the scattering vector is limited to $2|\vec{k}_i|$ (corresponding to the case of backscattering).

Let us consider the two lattice atoms, one at the origin and the other one at the position $\vec{R} = \sum m_i \vec{a}_i$ from the origin, where the $\vec{a}_i$ are the primitive vectors of the lattice and $m_i$ are integers (Fig. A.2). What is the difference in path length between a wave scattered at the origin and a wave that is scattered at position $\vec{R}$? Compared to the atom at the origin, the incoming wave front has to travel the additional path $R\cos\alpha = \vec{R} \cdot \vec{k}_i / k_i$ to reach the second atom. On the other hand, the distance between the second atom and the detector is shorter by $\vec{R} \cdot \vec{k}_f / k_f = \vec{R} \cdot \vec{k}_f / k_i$.

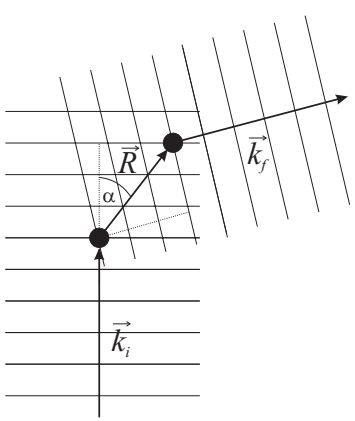



**Figure A.2:** Path difference between the scattered waves from two atoms with a relative distance $\vec{R}$

The total path difference for the two waves scattered in the direction of $\vec{k}_f$ by the two atoms is given by $\Delta = \frac{1}{k_i}\vec{R}\,(\vec{k}_i - \vec{k}_f) = \frac{1}{k_i}\vec{R}\,\vec{q}$. Constructive interference occurs for

$$\Delta = \frac{1}{k_i}\vec{q}\,\vec{R} = n\lambda = n\frac{2\pi}{k_i} \tag{A.4}$$

This means the scattering vector has to fulfill the condition

$$\vec{q}\,\vec{R} = 2\pi n \tag{A.5}$$

Equation (A.5) is known as the Laue[2] condition for constructive interference.

How are the Laue condition and the Bragg condition connected? In Fig. A.3 the wave vectors of the incident and outgoing radiation and the scattering vector are drawn for the Bragg reflection of Fig. A.1. We can conclude that for specular reflection, the scattering vector $\vec{q}$ is always perpendicular to the lattice plane. Its length is given by

$$|\vec{q}| = 2|\vec{k}_i| \sin\vartheta = \frac{4\pi}{\lambda} \cdot \sin\vartheta \tag{A.6}$$

To get constructive interference, we have to fulfill the Bragg condition. Inserting Eq. (A.1) into Eq. (A.6) leads to $|\vec{q}| = 2\pi/d$. In other words: we observe a diffraction peak, if the scattering vector is perpendicular to any of the lattice planes and its norm is equal to

$$|\vec{q}| = n \cdot \frac{2\pi}{d}, \tag{A.7}$$

with $d$ being the lattice spacing between this set of lattice planes.

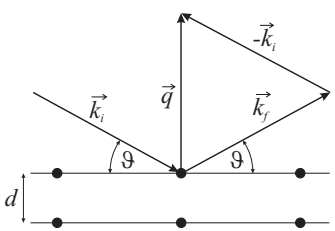



**Figure A.3:** Incident wave vector $\vec{k}_i$, outgoing wave vector $\vec{k}_f$, and scattering vector $\vec{q}$ for Bragg reflection (Fig. A.1) The length of the scattering vector is given by $|\vec{q}| = 2|k_i| \sin\theta = \frac{\pi}{\lambda} \sin\theta$.

### A.1.3 The reciprocal lattice

Let us come back to our task to find all possible diffraction peaks for a given crystal lattice. What are the possible scattering vectors that lead to constructive interference? This question can be answered in an elegant way by defining the so-called reciprocal lattice: If $\vec{a}_1$, $\vec{a}_2$, and $\vec{a}_3$ are primitive vectors of the crystal lattice, we choose a new set of vectors according to

$$\begin{aligned}
\vec{b}_1 &= 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \\
\vec{b}_2 &= 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \\
\vec{b}_3 &= 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}
\end{aligned} \tag{A.8}$$

[2] Max von Laue, 1879 – 1960. German professor for physics. Nobel prize for physics 1914 for the discovery of X-ray diffraction in crystals.

Just as a reminder: The dots between the vectors denote the scalar (inner) product and the crosses denote the cross (outer) product of the vectors. These vectors $\vec{b}_i$ are in units of $\mathrm{m}^{-1}$, which is proportional to the inverse of the lattice constants of the real space crystal lattice. This is why one calls the three-dimensional space spanned by these vectors the reciprocal space and the lattice defined by these primitive vectors is called the reciprocal lattice. These primitive reciprocal vectors have the following properties:

- $\vec{b}_1$ is perpendicular to $\vec{a}_2$ and $\vec{a}_3$ and therefore normal to the lattice plane spanned by $\vec{a}_2$ and $\vec{a}_3$. Accordingly, $\vec{b}_2$ is perpendicular to $\vec{a}_1$ and $\vec{a}_3$, and $\vec{b}_3$ is perpendicular to $\vec{a}_1$ and $\vec{a}_2$. (This is ensured by the cross product in the denominator.)
- The length of each vector $\vec{b}_j$ is $2\pi/d_j$, where $d_j$ is the distance between the the lattice planes perpendicular to $\vec{b}_j$. (This is ensured by the numerators $\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3$ in Eq. A.9.) Therefore, each of these primitive reciprocal lattice vectors fulfills the Laue condition Eq. (A.7) and is a possible scattering vector for constructive interference.

These two properties can be summarized into:

$$\vec{a}_j \cdot \vec{b}_k = 2\pi\delta_{jk} = \begin{cases} 0 & \text{for} \quad j \neq k \\ 2\pi & \text{for} \quad j = k \end{cases} \tag{A.9}$$

Another property is that for every set of parallel lattice planes, there are reciprocal lattice vectors that are normal to these planes. The shortest one of these reciprocal lattice vectors is used to characterize the plane orientation. The components $(h, k, l)$ of this vector are called Miller indices and the direction of the plane is denoted by $(hkl)$ for the single plane or $\{hkl\}$ for a set of planes (see Section 8.2.1).

Why is it useful to introduce such a complicated set of vectors? This becomes obvious when we look at the scalar product between a real space lattice vector $\vec{R}$ and a reciprocal lattice vector $\vec{q}$. Expressing these vectors by the corresponding primitive vectors we can write:

$$\begin{aligned} \vec{R} \cdot \vec{q} &= (l_1\vec{a}_1 + l_2\vec{a}_2 + l_3\vec{a}_3) \cdot (m_1\vec{b}_1 + m_2\vec{b}_2 + m_3\vec{b}_3) \\ &= l_1 m_1 \pi + l_2 m_2 \pi + l_3 m_3 \pi = n\pi \end{aligned} \tag{A.10}$$

with $n = l_1 m_1 + l_2 m_2 + l_3 m_3$ being an integer. This means that any reciprocal lattice vector automatically fulfills the Laue condition (Eq. A.5)! Therefore, all reciprocal lattice vectors are possible scattering vectors that lead to constructive interference. A more detailed analysis shows, that also every scattering vector that leads to constructive interference is a reciprocal lattice vector, i.e., the reciprocal lattice contains all possible scattering vectors More precisely: the vectors from the origin to every point of the reciprocal lattice give the set of possible scattering vectors with constructive interference for the corresponding real space lattice. For an infinite crystal this would lead to an infinite number of vectors. This raises the question, how many diffraction peaks we can expect to observe in a diffraction experiment. To answer this question, we have to remember that for elastic scattering, the length of the scattering vector cannot become larger than $2|\vec{k}_i|$. Taking into account this additional limit

$$|\vec{q}| \leq 2\,|\vec{k}_i|, \tag{A.11}$$

we end up with a small number of diffraction peaks.

### A.1.4 Ewald construction

The so-called Ewald[3] construction (Fig. A.4) allows us to find all possible scattering vectors $\vec{q}$ for a given incident wave vector $\vec{k}_i$ and a given crystal lattice. It is a simple geometric construction that automatically includes the additional boundary condition Eq. (A.11). As the first step, we construct the reciprocal lattice as described above. Take one of the lattice points as the origin. Draw the wave vector $\vec{k}_i$ of the incident wave starting from this origin. Then draw a sphere with radius $|\vec{k}_i|$ around the end point of $\vec{k}_i$ (so that it passes through the origin). If we then draw the outgoing wave vector $\vec{k}_f$ ending in the center of the sphere, its starting point will always be on the surface of the sphere. The scattering vector $\vec{q} = \vec{k}_f - \vec{k}_i$ then has to connect this starting point with the origin. For constructive interference to occur, the scattering vector $\vec{q}$ has to be a lattice vector of the reciprocal lattice (i.e. it must connect two lattice points). This is possible only if some reciprocal lattice points (in addition to the origin) lie on the surface of this sphere. Therefore, we get diffraction peaks only for very special values of $\vec{k}_i$ and $\vec{q}$. This means that usually we have to vary either the wavelength (this corresponds to vary the radius of the Ewald sphere) or the direction of the incident beam (this corresponds to a rotation of the sphere around the origin) to observe a reasonable number of the possible diffraction peaks.

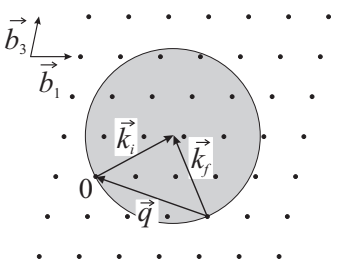



**Figure A.4:** The Ewald construction. Given an incident wave vector $\vec{k}_i$, a sphere of radius $|\vec{k}_i|$ is drawn around the end point of $\vec{k}_i$. Diffraction peaks are observed only if the scattering vector $\vec{q}$ ends on this sphere.

## A.2 Diffraction at Surfaces

The formalism of the reciprocal lattice and the Ewald construction can be applied to the diffraction at surfaces. As an example, we consider how the diffraction pattern of a LEED experiment (see Fig. 8.21) results from the surface structure. The most simple case is an experiment where the electron beam hits the crystal surface perpendicularly as shown in Fig. A.5. Since we do not have a Laue condition to fulfill in the direction normal to the surface, we get rods vertical to the surface instead of single points. All intersecting points between these rods and the Ewald sphere will lead to diffraction peaks. Therefore, we always observe diffraction

[3] Paul P. Ewald, 1888–1985. Physicist and crystallographer, professor at the Universities of Munich, Stuttgart, Queen's University at Belfast, and the Polytechnic Institute of Brooklyn.

peaks in LEED if the wave vector of the electrons is longer than the shortest reciprocal lattice vector, i.e., if the wavelength of the electrons is short enough. This is in contrast to the situation for three-dimensional diffraction, where diffraction peaks are observed only for special combinations of wave and scattering vectors. From Fig. A.5 we can see that the intersections between the rods and the Ewald sphere occur at the points where the component $\vec{q}_{||}$ of the scattering vector parallel to the surface is identical to a reciprocal lattice vector of the surface lattice. Therefore, the Laue condition for surface diffraction is given by:

$$\vec{q}_{||} \cdot \vec{R} = 2\pi n \qquad , \tag{A.12}$$

where $\vec{R}$ is a lattice vector of the real space surface lattice and $n$ is an integer. The number of diffraction peaks observed in a real experiment will be, however, limited by the fact that only radiation back scattered (i.e. with $\vec{k}_f$ directed away from the surface) will reach the detector.

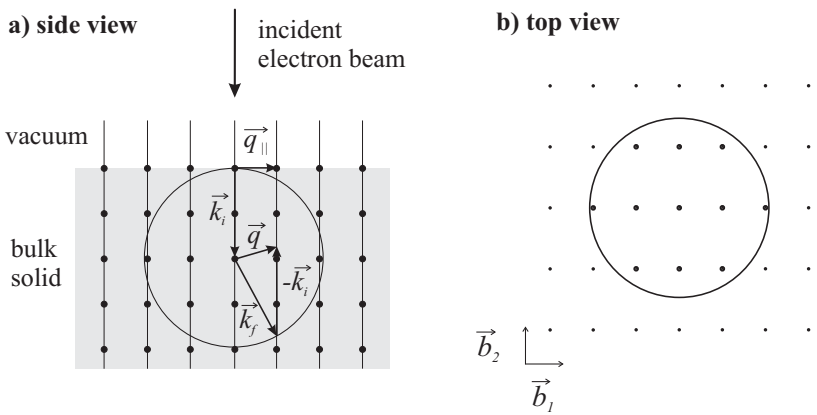



**Figure A.5:** Ewald construction for surface diffraction. a) a side view of the reciprocal lattice at the surface. Constructive interference occurs for all intersection points of the vertical rods with the Ewald sphere. This is equivalent to the condition when the component $\vec{q}_{||}$ of the scattering vector parallel to the surface is identical to a reciprocal lattice vector of the surface lattice. b) the top view of the reciprocal surface lattice. The circle is the projection of the Ewald sphere. If we disregarding the radiation scattered into the crystal, the number of lattice points within the circle (corresponding to the intersections of the rods with the Ewald sphere) is identical to the maximum number of observed diffraction peaks.

■ **Example A.1.** LEED pattern of the Cu(110) surface at an electron energy of 70 eV. Copper has an face centered cubic lattice with a lattice constant of 3.61 Å (Fig. A.6). The (110) surface has therefore a rectangular lattice with lattice constants $a_1 = 3.61$ Å and $a_2 = 3.61\,\text{Å} \cdot \sqrt{2}/2 = 2.55\,\text{Å}$ (Fig. 8.2). The reciprocal lattice is a rectangular lattice with the lattice constants $b_1 = 2\pi/3.61\text{Å} = 1.74\,\text{Å}^{-1}$ and $b_2 = 2\pi/2.55\,\text{Å} = 2.46\,\text{Å}^{-1}$. The wavelength of the electrons is (Eq. 8.11) $\lambda = h/\sqrt{2m_e eU} = 1.47\,\text{Å}$. The radius of the Ewald sphere is given by $|\vec{k}_i| = 2\pi/1.47\,\text{Å} = 4.27\,\text{Å}^{-1}$. If the electrons approach normal to the surface, the resulting pattern is therefore a rectangular lattice with relative lengths of the sides of $1.74 : 2.46$ and consisting of 15 diffraction spots.

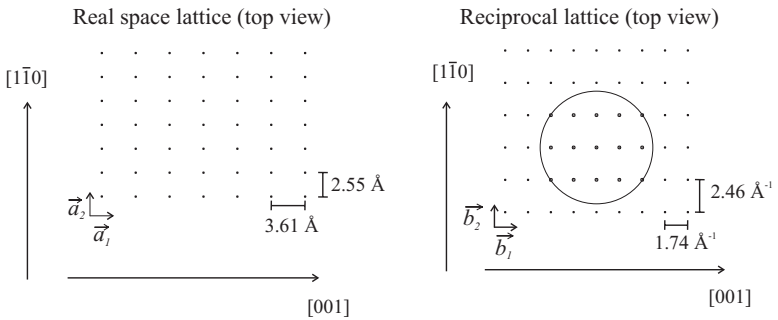



**Figure A.6:** Ewald construction for surface diffraction at the Cu(110) surface.

## A.3 Intensity of diffraction peaks

The Laue and the Bragg condition give us information about the angular distribution of the diffraction peaks. To calculate the peak intensities, we have to know more about the scattering properties of the atoms or molecules in the crystal. In the case of X-rays and electrons the scattering probability is proportional to the electron density $n_e(\vec{r})$ within the crystal. Since $n_e(\vec{r})$ has to have the same periodicity as the crystal lattice, we can write it as a three-dimensional Fourier series (using the notation $e^{ikx} = \cos kx + i \sin kx$):

$$\begin{aligned} n_e(x,y,z) &= \sum_{klm=-\infty}^{+\infty} n_{klm} e^{i(kx+ly+mz)} = \\ &= \sum_{klm=-\infty}^{+\infty} n_{klm} e^{i(k,l,m)\cdot(x,y,z)} = \\ &= \sum_{\vec{g}} n_{\vec{g}} e^{i\vec{g}\cdot\vec{r}} \end{aligned} \tag{A.13}$$

where $n_{\vec{g}}$ are the Fourier coefficients. From the requirement of lattice periodicity we can conclude $n_e(\vec{r}) = n_e(\vec{r}+\vec{R})$ for any lattice vector $\vec{R} = h\vec{a}_1 + k\vec{a}_2 + l\vec{a}_3$. Written as a Fourier series:

$$n_e(\vec{r}) = \sum_{\vec{g}} n_{\vec{g}} e^{i\vec{g}\cdot r+\vec{R}} = \sum_{g} n_g e^{i\vec{g}\cdot r} e^{i\vec{g}\cdot\vec{R}} \tag{A.14}$$

$$\begin{aligned} n_e(\vec{r}) &= n_e(\vec{r}) e^{i\vec{g}\cdot\vec{R}} \Rightarrow \\ 1 &= e^{i\vec{g}\cdot\vec{R}} = \cos\left(\vec{g}\cdot\vec{R}\right) + i \sin\left(\vec{g}\cdot\vec{R}\right) \Rightarrow \\ \vec{g}\cdot\vec{R} &= 2\pi n \end{aligned} \tag{A.15}$$

The last condition is identical to Eq. (A.10). Therefore, $\vec{g}$ must be a reciprocal lattice vector. This shows us another important property of the reciprocal lattice: it is the Fourier transform of the real space lattice.

The procedure of obtaining the crystal structure from a diffraction experiment should be straightforward. From the diffraction pattern we get the reciprocal lattice vectors and can construct the reciprocal lattice. The atomic arrangement should then be easily calculated by an

inverse Fourier transformation. Unfortunately, this is not possible for several reasons. The observed diffraction pattern contains only the intensity, but not the phase of the diffracted waves. The reconstruction is thus not unique. In addition, multiple diffraction events, which were not accounted for in the above calculations, might change the diffraction pattern. Therefore the analysis is done going the opposite way. First, a plausible atomic arrangement is assumed. From this, we calculate the diffraction pattern and compare it with the experimental results. The initial atomic arrangement is then adjusted until calculated and experimental diffraction patterns agree.

Up to this point we did not make any specific assumptions about the real space lattice. It could contain more than one atom per lattice point and more than more than one type of atoms. In such a case the lattice would be described using a Bravais lattice plus a basis (see Section 8.2.2. To obtain the intensity of the diffracted wave for crystals with a basis, we simply have to sum up the contributions from all scattering points within the unit cell. The scattering probability for a crystal of $N$ unit cells with an electron density $n_e(\vec{r})$ is proportional to:

$$N \int\limits_{\text{unit cell}} n_e(\vec{r})\, e^{-i\vec{q}\vec{r}}\, dV = NS_G, \tag{A.16}$$

where the integration runs over the volume of one unit cell. $S_G$ is the so-called geometrical structure factor. It contains information about the positions of the atoms within the unit cell and takes into account the interference between the rays scattered from the different basis sites within the unit cell. Since the intensity of a diffracted wave is proportional to the square of the amplitude, the measured intensities of the diffraction peaks are proportional to $|S_G|^2$.

If the basis contains atoms of different elements it is useful to separate the phase shifts due to the geometric arrangement from the scattering characteristics of the atoms. If $\vec{r}_j$ is the vector from the origin of the unit cell to the center of atom $j$ (denoting the position of atom $j$ within the unit cell), and $n_j(\vec{r}-\vec{r}_j)$ is the contribution of atom $j$ to the total electron density at $\vec{r}$, the structure factor can be written as:

$$\begin{aligned} S_G &= \sum_j \int n_j(\vec{r}-\vec{r}_j)\, e^{-i\vec{q}\vec{r}}\, dV \\ &= \sum_j e^{-i\vec{q}\vec{r}_j} \int n_j(\vec{\rho})\, e^{-i\vec{q}\vec{\rho}}\, dV \\ &= \sum_j F_j\, e^{-i\vec{q}\vec{r}_j} \end{aligned} \tag{A.17}$$

with $\vec{\rho} = \vec{r} - \vec{r}_j$ and the atomic form factor

$$F_j = \int n_j(\vec{\rho})\, e^{-i\vec{q}\cdot\vec{\rho}}\, dV \tag{A.18}$$

that is determined by the scattering characteristics of the atom $j$.
If we write $\vec{r}_j = x_j\vec{a}_1 + y_j\vec{a}_2 + z_j\vec{a}_3$ we get

$$\vec{q}\cdot\vec{r}_j = (x_j\vec{a}_1 + y_j\vec{a}_2 + z_j\vec{a}_3)\cdot(h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3) = 2\pi(x_jh + y_jk + z_jl) \tag{A.19}$$

and

$$S_G(hkl) = \sum_j F_j e^{-i2\pi(x_j h + y_j k + z_j l)} \tag{A.20}$$

as the structure factor of the (hkl) plane.

The atomic form factor accounts for the internal structure of the different atoms or molecules. It will also be different for X-rays and neutrons, since the former probe the electron distribution of the target, while the latter interact with the nuclei of the atoms. Therefore, the analysis of the positions of the reflexes indicates mainly the lattice constants and angles. The intensity of the reflexes contains mainly information about the atomic configuration within an unit cell (structure factor) and the scattering behavior of the single atoms (form factor).

The description of a given crystal by a lattice and a basis is not unique. One might e.g. choose to double the size of the unit cell by including more atoms into the basis. This would also lead to a different reciprocal lattice. This seams to lead to a contradiction, since the diffraction pattern should only depend on the crystal and not how we choose our description. As we will see in example A.2, the choice of a different basis leads to a change in the structure factor so that the combination of reciprocal lattice and structure factor always leads to the same diffraction pattern.

■ **Example A.2.** A body centered cubic lattice can also be regarded as a simple cubic lattice with primitive lattice vectors ($a$,0,0), (0,$a$,0) and (0,0,$a$), and a basis $\vec{r}_1 = 0$ and $\vec{r}_2 = \frac{a}{2}(1,1,1)$. The reciprocal lattice will then be a simple cubic lattice with primitive lattice vectors

$$\frac{2\pi}{a}(1,0,0), \qquad \frac{2\pi}{a}(0,1,0), \qquad \frac{2\pi}{a}(0,0,1) \tag{A.21}$$

and with a structure factor

$$S_G = F(1 + e^{-i\frac{a}{2}\vec{q}\cdot(1,1,1)}) \tag{A.22}$$

Substituting

$$\vec{q} = \frac{2\pi}{a} \cdot (n_1 \cdot (1,0,0) + n_2 \cdot (0,1,0) + n_3 \cdot (0,0,1)) \tag{A.23}$$

into Eq. (A.22) we get

$$\begin{aligned} S_G &= F(1 + e^{-i\pi(n_1+n_2+n_3)}) = F(1 + (-1)^{n_1+n_2+n_3}) \\ &= \begin{cases} 2F & \text{for} \quad n_1+n_2+n_3 \quad \text{even} \\ 0 & \text{for} \quad n_1+n_2+n_3 \quad \text{odd} \end{cases} \end{aligned} \tag{A.24}$$

with $n_1$, $n_2$, and $n_3$ being integers. Points that have an odd sum of coordinates in the simple cubic reciprocal lattice do not lead to diffraction peaks. This makes it equivalent to a reciprocal face centered cubic structure with side length $4\pi/a$, which is precisely the reciprocal lattice we would have obtained if we had treated the body centered cubic lattice as a bravais lattice rather than as a lattice with basis.

# B Symbols and abbreviations

Many symbols are not unique for a certain physical quantity but are used two or even three times. We use the symbols as they are usually used in the relevant literature. Since the scope of this book includes many disciplines and thus different scientific communities, multiple usage of symbols is unavoidable. In molecular chemistry and physics, for instance, $\mu$ is the dipole moment while in engineering, $\mu$ symbolizes the friction coefficient.

| | |
|---|---|
| $A$ | Area ($m^2$) |
| $A_H$ | Hamaker constant (J) |
| $a$ | Activity (mol $m^{-3}$), contact radius (m), frequency factor (Hz) |
| $\vec{a}_1, \vec{a}_2$ | Unit vectors of a two-dimensional unit cell of lengths $a_1$ and $a_2$ (m) |
| $a_h$ | Hydraulic radius (m) used to characterize pore sizes |
| $\vec{b}_1, \vec{b}_2$ | Unit vectors of a reconstructed unit cell of lengths $b_1$ and $b_2$ (m) or reciprocal lattice vectors |
| $C$ | Capacitance (C $V^{-1}$), van der Waals coefficient (J $m^6$), BET constant |
| $C^A$ | Differential capacitance per unit area (C $V^{-1}m^{-2}$) |
| $C_1$, $C_2$ | Two principal curvatures ($m^{-1}$) |
| $C_0$ | Spontaneous curvature ($m^{-1}$) |
| CMC | Critical micelle concentration (mol $L^{-1}$) |
| $c$ | Concentration or number density of molecules (number per $m^3$, mol $m^{-3}$, or mol $L^{-1}$= M) |
| $\vec{c}_1, \vec{c}_2$ | Unit vectors of a two-dimensional overlayer of lengths $c_1$ and $c_2$ (m) |
| $D$ | Distance (m), diffusion coefficient ($m^2s^{-1}$) |
| $D_0$ | Interatomic spacing used to calculate adhesion (m), typically 1.7 Å |
| $E$ | Electric field strength (V $m^{-1}$), Youngs modulus (Pa) |
| $E^*$ | Reduced Youngs modulus (Pa) |
| $E_F$ | Fermi energy (J) |
| $F$ | Helmholtz free energy (J), force (N) |
| $F_{adh}$ | Adhesion force (N) |
| $F_F$ | Friction force (N) |
| $F_L$ | Load (N) |
| $F^\sigma$, $f^\sigma$ | Interfacial free (Helmholtz) energy in Gibbs convention (J) and interfacial free (Helmholtz) energy per unit area (J $m^{-2}$), respectively |
| $f$ | Force per unit area (N $m^{-2}$) |
| $G$ | Gibbs free energy (J) |

*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl
Copyright © 2003 Wiley-VCH Verlag & Co. KGaA
ISBN: 3-527-40413-9

| | |
|---|---|
| $G_m,\ G_m^0$ | Molar Gibbs free energy and standard molar Gibbs free energy (J mol$^{-1}$) |
| $G^\sigma,\ g^\sigma$ | Interfacial Gibbs free energy (J) in Gibbs convention and interfacial Gibbs free energy per unit area (J m$^{-2}$), respectively |
| $H$ | Enthalpy (J) |
| $h$ | Height of a liquid with respect to a reference level (m), Planck constant, layer thickness (m) |
| $I$ | Electric current (A) |
| $J$ | Nucleation rate (s$^{-1}$ m$^{-3}$) |
| $l$ | Length of one chain link (m) in a polymer chain |
| $K$ | Spring constant (Nm$^{-1}$), equilibrium constant (e.g. mol/L) |
| $k$ | Bending rigidity (J) |
| $\bar{k}$ | Saddle-splay modulus (J) |
| $L_0$ | Thickness of a polymer brush (m) |
| $M$ | Molar mass (kg mol$^{-1}$), torque in Ch. 11 (Nm) |
| $m$ | Mass (kg), molecular mass (kg per molecule) |
| $N$ | Number of molecules (dimensionless or mol), number of segments in a linear polymer chain |
| $N_{agg}$ | Aggregation number of surfactant micelles |
| $N_i$ | Number of molecules of a certain species $i$ (dimensionless or mol) |
| $N_S$ | Surfactant parameter |
| $n$ | Refractive index, integer number |
| $n_C$ | Number of carbon atoms in an alkyl chain |
| $P$ | Pressure (Pa), probability |
| $P_0$ | Equilibrium vapor pressure of a vapor in contact with a liquid having a planar surface (Pa) |
| $P_0^K$ | Equilibrium vapor pressure of a vapor in contact with a liquid having a curved surface (Pa) |
| $P_m$ | Yield pressure (Pa) |
| $p$ | Momentum (kg m s$^{-1}$), integer coefficient, exponent in Freundlich adsorption isotherm |
| $Q$ | Electric charge (As), heat (J), quality factor of a resonator |
| $q$ | Heat per unit area (J m$^{-2}$), integer coefficient |
| $R$ | Radius of a (usually) spherical object (m), gas constant |
| $R_1, R_2$ | Two principal radii of curvature (m) |
| $R_b$ | Radius of a spherical bubble (m) |
| $R_d$ | Radius of a spherical drop (m) |
| $R_g$ | Radius of gyration of a polymer (m) |
| $R_p$ | Radius of a spherical particle (m) |
| $R_0$ | Size of a polymer chain (m) |
| $r$ | Radius (m), radial coordinate in cylindrical or spherical coordinates |
| $r_C$ | Radius of a capillary (m) |
| $S$ | Entropy (J K$^{-1}$), number of adsorption binding sites per unit area (mol m$^{-2}$), spreading coefficient (N m$^{-1}$) |
| $S_0, S_1$ | Number of vacant and occupied adsorption binding sites per unit area (mol m$^{-2}$) |

| | |
|---|---|
| $S^\sigma, s^\sigma$ | Interfacial entropy in Gibbs convention($\mathrm{JK}^{-1}$) and interfacial entropy per unit area ($\mathrm{JK}^{-1}\mathrm{m}^{-2}$), respectively |
| $T$ | Temperature (K) |
| $T_\Theta$ | Theta temperature (K) |
| $t$ | Time (s) |
| $U$ | Internal energy (J), applied or measured electric potential (V) |
| $U^\sigma$ | Interfacial internal energy (J) |
| $V$ | Volume ($\mathrm{m}^3$) |
| $V_m$ | Molar volume ($\mathrm{m}^3$ $\mathrm{mol}^{-1}$) |
| $v$ | Velocity ($\mathrm{m\ s}^{-1}$) |
| $W$ | Helmholtz free energy (J), work (J) |
| $w$ | Helmholtz free energy for the interaction between two surfaces per unit area ($\mathrm{J\ m}^{-2}$) |
| $x, y, z$ | Cartesian coordinates (m), $y$ is also the reduced electric potential |
| $Z$ | Valency of an ion |
| $\alpha$ | Polarizability ($\mathrm{C\ m}^2\ \mathrm{V}^{-1}$), angle, accommodation coefficient |
| $\gamma$ | Surface tension ($\mathrm{N\ m}^{-1}$). Specifically, $\gamma_L$ and $\gamma_{SL}$ are the surface tensions of a liquid–vapor and a solid–liquid interface, respectively |
| $\Gamma$ | Interfacial excess ($\mathrm{mol\ m}^{-2}$), grafting density of polymer (mol $\mathrm{m}^{-2}$ or $\mathrm{m}^{-2}$) |
| $\delta$ | Thickness of the hydration layer (m), indentation (m) |
| $\Delta$ | In connection with another symbol it is a difference, ellipsometric parameter |
| $\varepsilon$ | Dielectric permittivity |
| $\zeta$ | Zeta potential (V) |
| $\eta$ | Viscosity (Pa s) |
| $\Theta$ | Contact angle (deg) |
| $\theta$ | Coverage of a surface by adsorbed molecules |
| $\vartheta$ | Inverse of the Debye length ($\mathrm{m}^{-1}$), line tension (N) |
| $\kappa_e$ | Electrical conductivity ($\mathrm{A\ V}^{-1}\ \mathrm{m}^{-1}$) |
| $\lambda$ | Wavelength (m) |
| $\lambda_D$ | Debye length (m) |
| $\mu$ | Chemical potential ($\mathrm{J\ mol}^{-1}$), dipole moment (C m), friction coefficient |
| $\mu_k, \mu_s$ | Coefficient of kinetic and static friction, respectively |
| $\mu_r$ | Coefficient of rolling friction (m) |
| $\nu$ | Frequency (Hz) |
| $\pi$ | Film pressure ($\mathrm{N\ m}^{-1}$) |
| $\Pi$ | Disjoining pressure (Pa) |
| $\rho$ | Mass density ($\mathrm{kg\ m}^{-3}$), molecular density of a pure phase (molecules per $\mathrm{m}^3$). In contrast, the density of dissolved molecules is denoted by $c$. |
| $\rho_e$ | Electric charge density ($\mathrm{C\ m}^{-3}$) |
| $\Upsilon$ | Surface stress ($\mathrm{J\ m}^{-2}$) |
| $\Phi$ | Thermionic work function (J) |
| $\phi$ | Volume fraction |
| $\varphi$ | Galvani potential (V) |
| $\Sigma$ | Specific surface area of a powder or porous material ($\mathrm{m}^2\ \mathrm{kg}^{-1}$), parameter which characterizes the degree of coincidence at grain boundaries |

| | |
|---|---|
| $\sigma$ | Surface charge density (C $m^{-2}$) |
| $\sigma_A$ | Area occupied by a single molecules on a surface ($m^2$), head group area of a surfactant ($m^2$) |
| $\tau_c$ | Yield stress ($Nm^{-2}$) |
| $\tau_s$ | Slip time |
| $\psi$ | Electric Volta potential (V) |
| $\Psi$ | Ellipsometric parameter |

# Fundamental constants

| | | |
|---|---|---|
| Atomic mass unit | $u$ | $1.66054\times10^{-27}$ kg |
| Avogadro constant | $N_A$ | $6.02214\times10^{23}$ $mol^{-1}$ |
| Boltzmann constant | $k_B$ | $1.38066\times10^{-23}$ J $K^{-1}$ |
| Electron mass | $m_e$ | $9.10939\times10^{-31}$ kg |
| Elementary charge | $e$ | $1.60218\times10^{-19}$ C |
| Faraday constant | $F_A = eN_A$ | 96485.3 C $mol^{-1}$ |
| Gas constant | $R = k_B N_A$ | 8.31451 J $K^{-1}$ $mol^{-1}$ |
| Planck constant | $h$ | $6.62608\times10^{-34}$ J s |
| Speed of light in vacuum | $c$ | $2.99792\times10^{8}$ m $s^{-1}$ |
| Standard acceleration of free fall | $g$ | 9.80665 m $s^{-2}$ |
| Vacuum permittivity | $\varepsilon_0$ | $8.85419\times10^{-12}$ A s $V^{-1}m^{-1}$ |

# Conversion factors

1 eV = $1.60218\times10^{-19}$ J
1 dyne = $10^{-5}$ N
1 erg = $10^{-7}$ J
1 kcal = 4.184 kJ
1 torr = 133.322 Pa = 1.333 mbar
1 bar = $10^5$ Pa
1 poise (P) = 0.1 Pa s
1 Debye (D) = $3.336\times10^{-30}$ Cm
1 V = 1 J/As
0°C = 273.15 K
$k_BT/e = 25.69$ mV at 25°C

# Bibliography


[1] A.W. Adamson, A.P. Gast, *Physical Chemistry of Surfaces*, 6th ed., John Wiley & Sons, New York, **1997**.

[2] G. Brezesinski, H.-J. Mögel, *Grenzflächen und Kolloide*, Spektrum Akademischer Verlag, Heidelberg, **1993**.

[3] R.J. Hunter, *Foundations of Colloid Science*, 2nd ed., Oxford University Press, Oxford, **2001**.

[4] D.F. Evans, H. Wennerström, *The Colloidal Domain – Where Physics, Chemistry, Biology, and Technology Meet*, Wiley-VCH, New York, **1999**.

[5] H.-D. Dörfler, *Grenzflächen und kolloid-disperse Systeme*, Springer Verlag, Berlin, **2002**.

[6] M.J. Schwuger, *Lehrbuch der Grenzflächenchemie*, Georg Thieme Verlag, Stuttgart, **1996**.

[7] J. Lyklema, *Fundamentals of Interface and Colloid Science I: Fundamentals*, Academic Press, San Diego, **1991**.

[8] J. Lyklema, *Fundamentals of Interface and Colloid Science II: Solid-Liquid Interfaces*, Academic Press, San Diego, **1995**.

[9] J. Lyklema, *Fundamentals of Interface and Colloid Science III: Liquid-Fluid Interfaces*, Academic Press, San Diego, **2000**.

[10] C. Panayiotou, *Langmuir* **2002**, *18*, 8841.

[11] J. Chang, J. Han, L. Yang, R.L. Jaffe, D.Y. Yoon, *J. Chem. Phys.* **2001**, *115*, 2831.

[12] G.M. Barrow, *Physical Chemistry*, McGraw-Hill, Auckland, **1979**, p. 43; P.W. Atkins, *Physical Chemistry*, 4th ed., Oxford University Press, Oxford, **1990**, p. 733.

[13] T. Tate, *Phil. Mag.* **1864**, *27*, 176.

[14] L. du Noüy, *J. Gen. Physiol.* **1919**, *1*, 521.

[15] C. Huh, S.G. Mason, *Colloid Polym. Sci.* **1975**, *253*, 566.

[16] L. Wilhelmy, *Annalen der Physik und Chemie*, **1863**, *119*, 177.

[17] W. Thomson, *Proc. Roy. Soc. Edinburgh* **1870**, *7*, 63.

[18] L.R. Fisher, J.N. Israelachvili, *J. Colloid Interface Sci.* **1981**, *80*, 528.

[19] L.R. Fisher, R.A. Gamble, J. Middlehurst, *Nature* **1981**, *290*, 575.

[20] H.K. Christenson, *J. Colloid Interface Sci.* **1988**, *121*, 170.

[21] J.E. Curry, H.K. Christenson, *Langmuir* **1996**, *12*, 5729.

[22] N.L. Cross, R.G. Picknett, *Trans. Faraday Soc.* **1963**, *59*, 846.



*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl


Copyright © 2003 Wiley-VCH Verlag & Co. KGaA


ISBN: 3-527-40413-9

[23] H. Schubert, *Chemie-Ing.-Techn.* **1973**, *45*, 396.
[24] A. Fogden, L.R. White, *J. Colloid Interface Sci.* **1990**, *138*, 414.
[25] A. de Lazzer, M. Dreyer, H.J. Rath, *Langmuir* **1999**, *15*, 4551.
[26] M. Volmer, A. Weber, *Zeitschr. f. phys. Chem* **1926**, *119*, 277.
[27] R. Becker, W. Döring, *Annalen der Physik* **1935**, *24*, 719.
[28] V. Vohra, R.H. Heist, *J. Chem. Phys.* **1996**, *104*, 382.
[29] A. Bertelsmann, R.H. Heist, *Atmospheric Research* **1998**, *46*, 195.
[30] J. Wölk, R. Strey, C.H. Heath, B.E. Wyslouzil, *J. Chem. Phys. B* **2002**, *117*, 4954.
[31] R.H. Heist, H. He, *J. Phys. Chem. Ref. Data* **1994**, *23*, 781.
[32] A. Laaksonen, V. Talanquer, D.W. Oxtoby, *Ann. Rev. Phys. Chem.* **1995**, *46*, 489.
[33] D.W. Oxtoby, *J. Phys.: Condens. Matter* **1992**, *4*, 7627.
[34] D.W. Oxtoby, *Accounts of Chemical Research* **1998**, *31*, 91.
[35] J. Wölk, R. Strey, *J. Chem. Phys. B* **2002**, *105*, 11683.
[36] S.F. Jones, G.M. Evans, K.P. Galvin, *Adv. Colloid Interface Sci.* **1999**, *80*, 27.
[37] G. Liger-Belair, M. Vignes-Adler, C. Voisin, B. Robillard, P. Jeandet, *Langmuir* **2002**, *18*, 1294.

[38] E.A. Guggenheim, *Thermodynamics*, North Holland, Amsterdam, $4^{th}$ ed., **1959**, p. 46.
[39] R. Aveyard, D.A. Haydon, *An Introduction to the Principles of Surface Chemistry*, University Press, Cambridge, **1973**.
[40] R. Defay, I. Prigogine, A. Bellemans, D.H. Everett, *Surface Tension and Adsorption*, Longmans & Green, London, **1966**.
[41] M. Kahlweit, *Grundzüge der Physikalischen Chemie VII: Grenzflächenerscheinungen*, Steinkopff Verlag Darmstadt, **1981**.
[42] D.H. Everett, *Pure Appl. Chem.* **1972**, *31*, 578.
[43] R.V. Eötvös, *Wied. Ann.* **1886**, *27*, 456.
[44] W. Ramsay, J. Shields, *Phil. Trans. A* **1883**, *184*, 647.
[45] R.G. Linford, *Chem. Rev.* **1978**, *78*, 81.
[46] A.I. Rusanov, *Surface Science Reports* **1996**, *23,* 173.
[47] H.-J. Butt, R. Raiteri, *Surface Science Reports* **1999**, *87*, 1.
[48] G. Nilsson, *J. Phys. Chem.* **1957**, *61*, 1135.
[49] A.M. Posner, J.R. Anderson, A.E. Alexander, *J. Colloid Sci.* **1952**, *7*, 623.
[50] P.K. Weissenborn, R.J. Pugh, *Langmuir* **1995**, *11*, 1422.
[51] R. Vuilleumier, V. Ego, L. Neltner, A.M. Cazabat, *Langmuir* **1995**, *11*, 4117.
[52] A.E. Hosoi, J.W.M. Bush, *J. Fluid Mech.* **2001**, *442*, 217.
[53] C. Marangoni, *Annalen der Physik und Chemie* **1871**, *143*, 337.

[54] G. Gouy, *Journ. de Phys.* **1910**, *9*, 457.
[55] D.L. Chapman, *Phil. Mag.* **1913**, *25*, 475.
[56] G. Gouy, *Ann. Phys.* **1917**, *7*, 129.
[57] P. Debye, E. Hückel, *Physikalische Zeitschrift* **1923**, *24*, 185.

[58] R. Raiteri, M. Grattarola, H.-J. Butt, *J. Phys. Chem.* **1996**, *100*, 16700.
[59] D.C. Grahame, *Chem. Rev.* **1947**, *41*, 441.
[60] G. Cevc, *Biochim. Biophys. Acta* **1990**, *1031-3,* 311.
[61] S.L. Carnie, G.M. Torrie, *Adv. Chem. Phys.* **1984**, *56*, 142.
[62] L. Blum, *Adv. Chem. Phys.* **1990**, *78*, 171.
[63] M. Queseda-Pérez, E. González-Tovar, A. Martin-Molina, M.Lozada-Cassou, R. Hidalgo-Alvarez, *Chem. Phys. Chem.* **2003**, *4*, 235.
[64] O. Stern, *Zeitschrift für Elektrochemie* **1924**, *30*, 508.
[65] R. Guidelli, W. Schmickler, *Electrochimica Acta* **2000**, *45*, 2317.
[66] E.J.W. Verwey, J.T.G. Overbeek, *Theory of the Stability of Lyophobic Colloids*, Elsevier, New York, **1948**, p. 51.
[67] D.Y.C. Chan, D.J. Mitchell, *J. Colloid Interface Sci.* **1983**, *95*, 193.

[68] M.G. Lippmann, *Ann. Chim. Phys.* **1875**, *5*, 494.
[69] A. Frumkin, O. Petry, B. Damaskin, *Electroanal. Chem. Interfacial Electrochem.* **1970**, *27*, 81.
[70] D.M. Mohilner, T. Kakiuchi, *J. Electrochem. Soc.* **1981**, *128*, 350.
[71] R.A. Fredlein, J. O'M. Bockris, *Surface Science* **1974**, *46*, 641.
[72] R. Raiteri, H.-J. Butt, *J. Phys. Chem.* **1995**, *99*, 15728.
[73] T.R. Beck, *J. Phys. Chem.* **1969**, *73*, 466.
[74] K.F. Lin, T.R. Beck, *J. Electrochem. Soc.* **1976**, *123*, 1145.
[75] I. Morcos, H. Fischer, *J. Electroanal. Chem*. **1968**, *17*, 7.
[76] O.J. Murphy, J.S. Wainright, *Langmuir* **1989**, *5*, 519.
[77] J.A.M. Sondag-Huethorst, L.G.J. Fokkink, *Langmuir* **1992**, *8*, 2560.
[78] D.M. Mohilner, T.R. Beck, *J. Phys. Chem.* **1979**, *83*, 1160.
[79] G. Lang, K.E. Heusler, *J. Electroanal. Chem.* **1994**, *377*, 1.
[80] W. Schmickler, E. Leiva, *J. Electroanal. Chem.* **1998**, *453*, 61.
[81] W. Paik, M.A. Genshaw, J.O'M. Bockris, *J. Phys. Chem*. **1970**, *74*, 4266.
[82] J. Wang, B.M. Ocko, A.J. Davenport, H.S. Isaacs, *Phys. Rev. B* **1992**, *46*, 10321.
[83] L. Perera, M.L. Berkowitz, *J. Phys. Chem.* **1993**, *97*, 13803.
[84] B.H. Bijsterbosch, J. Lyklema, *Adv. Colloid Interface Sci.* **1978**, *9*, 147.
[85] J.S. Noh, J.A. Schwarz, *J. Colloid Interface Sci.* **1989**, *130*, 157.
[86] C.P. Huang, W. Stumm, *J. Colloid Interface Sci.* **1973**, *43*, 409.
[87] D.E. Yates, S. Levine, T.W. Healy, *J. Chem. Soc. Faraday Trans. I* **1974**, *70*, 1807.
[88] J.A. Davies, J.O. Lecki, *J. Colloid Interface Sci.* **1980**, *74*, 32.
[89] R. Charmas, W. Piasecki, W. Rudzinski, *Langmuir* **1995**, *11*, 3199.
[90] D.A. Sverjensky, N. Sahai, *Geochimica et Cosmochimica Acta* **1996**, *60*, 3773.
[91] M. Kosmulski *J. Colloid Interface Sci.* **2002**, *253*, 77.
[92] L.S. Cerovic, S.K. Milonjic, D. Bahloul-Houlier, B. Doucey, *Colloids Surf. A* **2002**, *197*, 147.

[93] K. Jasmund, G. Lagaly (eds.) *Tonminerale und Tone. Struktur, Eigenschaften, Anwendungen und Einsatz in Industrie und Umwelt*, Steinkopff Verlag, Darmstadt, **1993**.

[94] J.S. Lyons, D.N. Furlong, T.W. Healy, *Aus. J. Chem.* **1981**, *34*, 1177.

[95] P.M. Claesson, P. Herder, P. Stenius, J.C. Eriksson, R.M. Pashley, *J. Colloid Interface Sci.* **1986**, *109*, 31.

[96] S. Nishimura, H. Tateyama, K. Tsunematsu, K. Jinnai, *J. Colloid Interface Sci.* **1992**, *152*, 359.

[97] P.J. Scales, T.W. Healy, D.F. Evans, *J. Colloid Interface Sci.* **1988**, *124*, 391.

[98] J. Lyklema, *J. Electroanal. Chem.* **1972**, *37*, 53.

[99] K. Furusawa, W. Norde, J. Lyklema, *Kolloid Z. u. Z. Polymere* **1972**, *250*, 908.

[100] W.P. Gomes, D. Vanmaekelbergh, *Electrochimica Acta* **1996**, *41*, 967.

[101] R.J. Hunter, *Zeta Potential in Colloid Science, Principles and Applications*, Academic Press, London, **1981**.

[102] H.J. Jakobasch; F. Simon, C. Werner, C. Bellmann, *Technisches Messen* **1996**, *63*, 439.

[103] M. Smoluchowski, in: *Handbuch der Electrizität und des Magnetismus Vol. II*, S. 366, Barth, Leipzig, **1921**.

[104] C. Werner, H. Körber, R. Zimmermann, S. Dukhin, H.J. Jacobasch, *J. Colloid Interface Sci.* **1998**, *208*, 329.

[105] D.C. Henry, *Proc. Roy. Soc. London A* **1931**, *133*, 106.

[106] A.S. Dukhin, T.G.M. van den Ven, *J. Colloid Interface Sci.* **1994**, *165*, 9.

[107] K. Wandelt, *Appl. Surf. Sci.* **1997**, *111*, 1.

[108] D. Cahen, A. Kahn, *Adv. Materials* **2003**, *15*, 271.

[109] J.N. Israelachvili, *Intermolecular and Surface Forces*, $2^{nd}$ ed., Academic Press, Amsterdam, **1992**.

[110] H. Margenau, N.R. Kestner, *Theory of Intermolecular Forces*, Pergamon Press, Oxford, **1969**.

[111] W.H. Keesom, *Physik. Zeitschr.* **1921**, *22*, 129+364.

[112] P. Debye, *Physik. Zeitschr.* **1920**, *21*, 178; **1921**, *22*, 302.

[113] F. London, *Zeitschr. f. Physik* **1930**, *63*, 245.

[114] J. Mahanty, B.W. Ninham, *Dispersion Forces*, Academic Press, New York, **1976**.

[115] H.C. Hamaker, *Physica* **1937**, *4*, 1058.

[116] D. Henderson, D.M. Duh, X. Chu, D. Wasan, *J. Colloid Interface Sci.* **1997**, *185*, 265.

[117] K. Autumn et al. *Proc. Natl. Acad. Sci. USA* **2002**, 192252799.

[118] E.M. Lifshitz, *Soviet Phys. JETP (Engl. Transl.)* **1956**, *2*, 73.

[119] I.E. Dzyaloshinskii, E.M. Lifshitz, L.P. Pitaevskii, *Adv. Phys.* **1961**, *10*, 165.

[120] J.N. Israelachvili, *Quart. Rev. Biophys.* **1974**, *6*, 341.

[121] A. Meurk, P.W. Luckham, L. Bergström, *Langmuir* **1997**, *13*, 3896.

[122] L.B. Boinovich, *Adv. Colloid Interf. Sci.* **1992**, *37*, 177.

[123] H.J. Schulze, *Physico-chemical Elementary Processes in Flotation*. Elsevier, Amsterdam, **1984**.

[124] V.A. Parsegian, G.H. Weiss, *J. Colloid Interf Sci.* **1981**, *81,* 285.
[125] J. Visser, *Adv. Colloid Interface Sci.* **1972**, *3,* 331.
[126] H.D. Ackler, R.H. French, Y.M. Chiang, *J. Colloid Interface Sci.* **1996**, *179*, 460.
[127] C.J. Drummond, G. Georgaklis, D.Y.C. Chan, *Langmuir* **1996**, *12*, 2617.
[128] C.J. Drummond, D.Y.C. Chan, *Langmuir* **1996**, *12*, 3356.
[129] L. Bergström, *Adv. Colloid Interface Sci.* **1997**, *70*, 125.
[130] C.M. Roth, B.L. Neal, A.M. Lenhoff, *Biophys. J.* **1996**, *70*, 977.
[131] J.N. Israelachvili, *Langmuir* **1994**, *10*, 3369.
[132] B.V. Derjaguin, *Kolloid Zeitschrift* **1934**, *69*, 155.
[133] L.R. White, *J. Colloid Interface Sci.* **1983**, *95*, 286.
[134] B.V. Derjaguin, N.V. Churaev, V.M. Muller, *Surface Forces*, Consultants Bureau, New York, **1987**.
[135] P.M. Claesson, T. Ederth, V. Bergeron, M.W. Rutland, *Adv. Colloid Interface Sci.* **1996**, *67*, 119.
[136] B.V. Derjaguin, A.S. Titijevskaia, I.I. Abricossova, A.D. Malkina, *Discuss. Faraday Soc.* **1954**, *18*, 24.
[137] G.C.J. Rouweler, J.T.G. Oberbeek, *Trans. Faraday Soc.* **1971**, *67*, 2117.
[138] W. Black, J.G.V. de Jongh, J.T.G. Overbeek, M.J. Sparnaay, *Trans. Faraday Soc.* **1960**, *56*, 1597.
[139] B.V. Derjaguin, Y.I. Rabinovich, N.V. Churaev, *Nature* **1978**, 313.
[140] D. Tabor, F.R.S. Winterton, R.H.S. Winterton, *Proc. Roy. Soc. London A* **1969**, *312*, 435.
[141] J.N. Israelachvili, D. Tabor, *Proc. Roy. Soc. London A* **1972**, *31*, 19.
[142] J.N. Israelachvili, G.E. Adams, *J. Chem. Soc. Faraday Trans. I* **1978**, *74*, 975.
[143] G. Binnig, C.F. Quate, C. Gerber, *Phys. Rev. Lett.* **1986**, *56*, 930.
[144] B. Cappella, G. Dietler, *Surf. Sci. Rep.* **1999**, *34*, 1.
[145] W.A. Ducker, T.J. Senden, R.M. Pashley, *Nature* **1991**, *353*, 239.
[146] H.-J. Butt, *Biophys. J.* **1991**, *60*, 1438.
[147] D.M. LeNeveu, R.P. Rand, V.A. Parsegian, *Nature* **1976**, *256*, 601.
[148] V.A. Parsegian, R.P. Rand, N.L. Fuller, D.C. Rau, *Methods in Enzymology* **1986**, *127*, 400.
[149] S. Leikin, V.A. Parsegian, D.C. Rau, R.P. Rand, *Ann. Rev. Phys. Chem.* **1993**, *44*, 369.
[150] G. Reiter, *Phys. Rev. Lett.* **1992**, *68*, 75.
[151] R. Seemann, S. Herminghaus, K. Jacobs, *Phys. Rev. Lett.* **2001**, *86*, 5534.
[152] R. Hogg, T.W. Healy, D.W. Fuerstenau, *Trans. Faraday Soc.* **1966**, *62*, 1638.
[153] V.A. Parsegian, D. Gingell, *Biophys. J.* **1972**, *12*, 1192.
[154] H. Schulze, *J. prakt. Chemie* **1882**, *25*, 431 and **1883**, *27*, 320.
[155] W.B. Hardy, *Proc. Roy. Soc. London* **1990**, *66*, 110; *Zeitschr. Physikal. Chemie* **1990**, *33*, 385.
[156] J.T.G. Overbeek, in: *Colloid Science, Vol. 1*, H.R. Kruyt (ed.), Elsevier, Amsterdam, **1952**, p. 278.

[157] B. Derjaguin, *Acta Physicochim. URSS* **1939**, *10*, 333.
[158] L.D. Landau, *Acta Physicochim. URSS* **1941**, *14*, 633.
[159] I. Langmuir, *J. Chem. Phys.* **1938**, *6*, 873.
[160] R.K. Schofield, *Trans. Faraday Soc.* **1946**, *42*, 219.
[161] J.B. Hemwall, P.F. Low, *Soil Sci.* **1956**, *82*, 135.
[162] L.A. Rowley, D. Nicholson, N.G. Parsonage, *Mol. Phys.* **1976**, *31*, 365.
[163] F.F. Abraham, *J. Chem. Phys.* **1978**, *68*, 3713.
[164] D.Y.C. Chan, D.J. Mitchell, B.W. Ninham, A. Pailthorpe, *Mol. Phys.* **1978**, *35*, 1669.
[165] I.K. Snook, D. Henderson, *J. Chem. Phys.* **1978**, *68*, 2134.
[166] A.K. Doerr, M. Tolan, T. Seydel, W. Press, *Physica B* **1998**, *248*, 263.
[167] C.J. Yu, A.G. Richter, J. Kmetko, A. Datta, P. Dutta, *Europhys. Lett.* **2000**, *50*, 487.
[168] W. van Megen, I. Snook, *J. Chem. Soc. Faraday Trans. II* **1979**, *7*, 1095.
[169] J.E. Lane, T.H. Spurling, *Chem. Phys. Lett.* **1979**, *67*, 107.
[170] P. Tarazona, L. Vicente, *Mol. Phys.* **1985**, *56*, 557.
[171] V.S.Mitlin, M.M. Sharma, *J. Colloid Interface Sci.* **1995**, *170*, 407.
[172] R.G. Horn, J.N. Israelachvili, *J. Chem. Phys.* **1981**, *75*, 1400.
[173] H.K. Christenson, *J. Chem. Phys.* **1983**, *78*, 6906.
[174] V. Franz, H.-J. Butt, *J. Phys. Chem. B* **2002**, *106*, 1703.
[175] J. Israelachvili, H. Wennerström, *Nature* **1996**, *379*, 219.
[176] H.-J. Butt, V. Franz, in: *Water in Biomaterials Surface Science*, M. Morra (ed.), John Wiley & Sons, Chichester, **2001**.
[177] J.N. Israelachvili, R.M. Pashley, *Nature* **1983**, *306*, 249.
[178] S. Marcelja, N. Radic, *Chem. Phys. Lett.* **1976**, *42*, 129.
[179] S. Marcelja, J.N. Israelachvili, H. Wennerström, *Nature* **1997**, *385*, 689.
[180] R. Lipowski, S. Grotehans, *Biophys. Chem.* **1994**, *49*, 27.
[181] L. Perera, U. Essmann, M.L. Berkowitz, *Langmuir* **1996**, *12*, 2625.
[182] J.N. Israelachvili, H. Wennerström, *J. Phys. Chem.* **1992**, *96*, 520.
[183] G. Cevc, M. Hauser, A.A. Kornyshev, *Langmuir* **1995**, *11*, 3103.
[184] H.K. Christenson, in: *Modern Approaches to Wettability*, M.E. Schrader, G. Loeb (eds.), Plenum Press, New York, **1992**, p. 29.
[185] J.N. Israelachvili, R. Pashley, *Nature* **1982**, *300*, 341.
[186] J.N. Israelachvili, R. Pashley, *J. Colloid Interface Sci*. **1984**, *98*, 500.
[187] P.M. Claesson, C.E. Blom, P.C. Herder, B.W. Ninham, *J. Colloid Interface Sci.* **2002**, *96*, 1.
[188] R. Steitz, T. Gutberlet, T. Hauss, B. Klösgen, R. Krastev, S. Schemmel, A.C. Simonsen, G.H. Findenegg, *Langmuir* **2003**, *19*, 2409.
[189] H.K. Christenson, P.M. Claesson, *Adv. Colloid Interface Sci.***2001**, *91*, 391.
[190] G. Strobel, *The Physics of Polymers*, $2^{nd}$ ed., Springer, Berlin, 1996.
[191] J. Klein, P.F. Luckham, *Macromolecules* **1984**, *17*, 1041.
[192] J.N. Israelachvili, M. Tirrel, J. Klein, Y. Almog, *Macromolecules* **1984**, *17*, 204.
[193] J. Marra, H.K. Christenson, *J. Phys. Chem.* **1989**, *93*, 7180.

[194] R.G. Horn, S.J. Hirz, G. Hadziioannou, C.W. Frank, *J. Chem. Phys.* **1989**, *90*, 6767.
[195] S.J. O'Shea, M.E. Welland, T. Rayment, *Langmuir* **1993**, *9*, 1826.
[196] A.S. Lea, J.D. Andrade, V. Hlady, *Colloids & Surfaces A* **1994**, *93*, 349.
[197] X. Chatellier, T.J. Senden, J.F. Joanny, J.M. di Meglio, *Europhys. Lett.* **1998**, *41*, 303.
[198] S.S. Patel, M.Tirrell, *Ann. Rev. Phys. Chem.* **1989**, *40*, 597.
[199] V. Bergeron, P.M. Claesson, *Adv. Colloid Interface Sci.* **2002**, *96*, 1.
[200] A.K. Dolan, S.F. Edwards, *Proc. Roy. Soc. London A* **1974**, *337*, 509.
[201] P.G. de Gennes, *Adv. Colloid Interface Sci.* **1987**, *27*, 189.
[202] S.T. Milner, T.A. Witten, M.E. Cates, *Macromolecules* **1988**, *21*, 2610.
[203] H.-J. Butt, M. Kappl, H. Müller, R. Raiteri, W. Meyer, J. Rühe, *Langmuir*, **1999**, *15*, 2559.
[204] H. Hertz, *J. Reine Angewandte Mathematik* **1882**, *92*, 156.
[205] K.L. Johnson, K. Kendall, A.D. Roberts, *Proc. Roy. Soc. London A* **1971**, *324*, 301.
[206] B.V. Derjaguin, V.M. Muller, Y.P. Toporov, *J. Colloid Interface Sci.* **1975**, *53*, 314.
[207] V.M. Muller, V.S. Yushchenko, B.V. Derjaguin, *J. Colloid Interface Sci.* **1983**, *92*, 92.
[208] D. Maugis, *J. Colloid Interface Sci.* **1992**, *150*, 243.
[209] M.D. Pashley, J.B. Pethica, D. Tabor, *Wear* **1984**, *100*, 7.
[210] H. Krupp, *Adv. Colloid Interface Sci.* **1967**, *1*, 111.
[211] R.I. Larsen, *Amer. Ind. Hyg. Assoc. J.* **1958**, *19*, 256.
[212] A.D. Zimon, *Kolloid Zh.* **1963**, *25,* 317.
[213] G. Böhme, W. Kling, H. Krupp, H. Lange, G. Sandstede, *Zeitschr. Angew. Physik* **1964**, *16*, 486.
[214] S.W. Booth, J.M. Newton, *J. Pharm. Pharmacol.* **1987**, *39,* 679.
[215] F. Podczeck, J.M. Newton, J.M. *J. Pharmaceutical Sci.* **1995**, *84*, 1067.
[216] M. Kappl, H.-J. Butt, *Part. Part. Syst. Charact.* **2002**, *19*, 129.
[217] J.N. Israelachvili, E. Perez, R.K. Tandon, *J. Colloid Interface Sci.* **1980**, *78*, 260.
[218] R.G. Horn, J.N. Israelachvili, F. Pribac, *J. Colloid Interface Sci.* **1987**, *115*, 480.
[219] L.O. Heim, J. Blum, M. Preuss, H.-J. Butt, *Phys. Rev. Lett.* **1999**, *83*, 3328.

[220] P.G. de Gennes, *Rev. Mod. Phy.* **1985**, *57*, 827.
[221] A.W. Neumann, J.K. Spelt, *Applied Surface Thermodynamics*, Surfactant Science Series 63, Marcel Dekker, New York, **1996**.
[222] T. Young, *Phil. Trans. Roy. Soc. London* **1805**, *95*, 65.
[223] P.S. de Laplace, *Mécanique Céleste*, suppl. au X. Livre, Croucier, Paris, 1805.
[224] B.A. Pethica, *J. Colloid Interface Sci.* **1977**, *62*, 567.
[225] J. Drelich, *Colloids & Surfaces A* **1996**, *116*, 43.
[226] T. Pompe, S. Herminghaus, *Phys. Rev. Lett.*, **2000**, *85*, 1930.
[227] V.G. Babak, *Colloids & Surfaces A* **1999**, *156*, 423.
[228] U. Zimmermann, H. Schneider, L.H. Wegner, H.J. Wagner, M. Szimtenings, A. Haase, F.W. Bentrup, *Physiologia Plantarum* **2002**, *114*, 327.

[229] A. Scheludko, B.V. Toshev, D.T. Bojadjiev, *J. Chem. Soc. Faraday Trans. I* **1976**, *72*, 2815.
[230] A.B.D. Cassie, S. Baxter, *Trans. Faraday Soc.* **1944**, *40*, 546.
[231] A.S. Dimitrov, P.A. Kralchevsky, A.D. Nikolov, H. Noshi, M. Matsumoto, *J. Colloid Interface Sci.* **1991**, *145*, 279.
[232] J.R. Dann, *J. Colloid Interface Sci.* **1970**, *32*, 302.
[233] F.E. Bartell, H.J. Osterhof, *Zeitschr. Physik. Chemie* **1927**, *130*, 715.
[234] H.G. Bruil, J. van Aartsen, *J. Colloid & Polymer Sci.* **1974**, *252*, 32.
[235] R. Lucas, *Kolloid-Z.* **1918**, *23*, 15.
[236] E.W. Washburn, *Phys. Rev.* **1921**, *17*, 273.
[237] S. Levine, J. Lowndes, E.J. Watson, G. Neale, *J. Colloid Interface Sci.* **1980**, *73*, 136.
[238] K. Grundke, A. Augsburg, *J. Adhesion Sci. Technol.***2000**, *14*, 765.
[239] J. van Brackel, *Powder Technol.* **1975**, *11*, 205.
[240] G. Martic, F. Gertner, D. Seveno, D. Coulon, J. De Coninck, T.D. Blake, *Langmuir* **2002**, *18*, 7971.
[241] S. Ecke, M. Preuss, H.-J. Butt, *J. Adh. Sci. Technol.* **1999**, *13*, 1181.
[242] R. Shuttleworth, G.L.J. Bailey, *Disc. Faraday Soc.* **1948**, *3*, 16.
[243] J.D. Eick, R.J. Good, A.W. Neumann, *J. Colloid Interface Sci*. **1975**, *53*, 235.
[244] C. Huh, S.G. Mason, *J. Colloid Interface Sci.* **1977**, *60*, 11.
[245] S. Brandon, A. Marmur, *J. Colloid Interface Sci.* **1996**, *183*, 351.
[246] J. Drelich, J.D. Miller, A. Kumar, G.M. Whitesides, *Colloids & Surfaces A* **1994**, *93*, 1.
[247] P.S. Swain, R. Lipowsky, *Langmuir* **1998**, *14*, 6772.
[248] A.W. Neumann, R.J. Good, *J. Colloid Interface Sci.* **1972**, *38*, 341.
[249] J.F. Joanny, P.G. de Gennes, *J. Chem. Phys.* **1984**, *81*, 552.
[250] M.E.R. Shanahan, A. Carré, *Langmuir* **1995**, *11*, 1396.
[251] R. Kern, P. Müller, *Surface Sci.* **1992**, *264*, 467.
[252] C.W. Extrand, Y. Kumagai, *J. Colloid Interface Sci.* **1996**, *184*, 191.
[253] A. Carré, J.C. Gastel, M.E.R. Shanahan, *Nature* **1996**, *379*, 432.
[254] X.W. Yang, *Appl. Phys. Lett.* **1995**, *67*, 2249.
[255] C.W. Extrand, *J. Colloid Interface Sci.* **1998**, *207*, 11.
[256] J. Drelich, J.D. Miller, R.J. Good, *J. Colloid Interface Sci.* **1996**, *179*, 37.
[257] E.L. Decker, S. Garoff, *Langmuir* **1997**, *13*, 6321.
[258] L.W. Schwartz, S. Garoff, *Langmuir* **1985**, *1*, 219.
[259] G. Wiegand, T. Jaworek, G. Wegner, E. Sackmann, *J. Colloid Interface Sci.* **1997**, *196*, 299.
[260] R.N. Wenzel, *Industrial and Engineering Chemistry* **1936**, *28*, 988.
[261] D. Quéré, *Physica A* **2002**, *313*, 32.
[262] W. Barthlott, C. Neinhuis, *Planta* **1997**, *202*, 1.
[263] T. Onda, S. Shibuichi, N. Satoh, K. Tsujii, *Langmuir* **1996**, *12*, 2125.
[264] W. Chen, A.Y. Fadeev, M.C. Hsieh, D. Öner, J. Youngblood, T.J. McCarthy, *Langmuir* **1999**, *15*, 3395.

[265] Z.Z. Gu, H. Uetsuka, K. Takahashi, R. Nakajima, H. Onishi, A. Fujushima, O. Sato, *Angew. Chem. Int. Ed.* **2003**, *42*, 894.

[266] A.B.D. Cassie, *Discuss. Faraday Soc.* **1948**, *3*, 11.

[267] E. Chibowksi, R. Perea-Carpio, *Adv. Colloid Interface Sci.* **2002**, *98*, 245.

[268] F. Brochard-Wyart, J.M. de Meglio, D. Quéré, P.G. de Gennes, *Langmuir* **1991**, *7*, 335.

[269] E.K. Yeh, J. Newman, C.J. Radke, *Colloids & Surfaces A* **1999**, *156*, 525.

[270] R.J. Good, L.A. Girifalco, *J. Phys. Chem.* **1960**, *64*, 561.

[271] F.M. Fowkes, *Ind. Eng. Chem.* **1964**, *56*, 40.

[272] D.K. Owens, R.D. Wendt, *J. Appl. Polym. Sci.* **1969**, *13*, 1741.

[273] J.C. Berg, in: *Wettability*, J.C. Berg (ed.), Marcel Dekker, New York, **1993**, p. 76.

[274] J.H. Clint, *Current Opinion in Colloid & Interface Sci.* **2001**, *6*, 28.

[275] T.D. Blake, in: *Wettability*, J.C. Berg (ed.), Marcel Dekker, New York, **1993**, p. 252.

[276] S.F. Kistler, in: *Wettability*, J.C. Berg (ed.), Marcel Dekker, New York, **1993**, p. 311.

[277] R.L. Hoffman, *J. Colloid Interface Sci.* **1975**, *50*, 228.

[278] T.D. Blake, Y.D. Shikhmurzaev, *J. Colloid Interface Sci.* **2002**, *253*, 196.

[279] W. Radigan, H. Ghiradella, H.L. Frisch, H. Schonhorn, T.K. Kwei, *J. Colloid Interface Sci.* **1974**, *49*, 241.

[280] N.V. Churaev, *Rev. Phys. Appl.***1988**, *23*, 975.

[281] L. Leger, M. Erman, A.M. Guinet-Picard, D. Ausserre, C. Strazielle, *Phys. Rev. Lett.* **1988**, *60*, 2390.

[282] I. Ispolatov, B. Widom, *Physica A* **2000**, *297*, 203.

[283] J. De Coninck, S. Hoorelbeke, M.P. Valignat, A.M. Cazabat, *Phys. Rev. E* **1993**, *48*, 4549.

[284] E. Bonaccurso et al., *Langmuir* **2002**, *18*, 8056.

[285] J. Yang, J. Koplik, J.R. Banavar, *Phys. Rev. Lett.* **1991**, *67*, 3539.

[286] J.A. Nieminen, D.B. Abraham, M. Karttunen, K. Kaski, *Phys. Rev. Lett*. **1992**, *69*, 124.

[287] A.M. Cazabat, M.P. Valignat, S. Villette, J. De Coninck, F. Louche, *Langmuir* **1997**, *13*, 4754.

[288] M. Steinhart, J.H. Wendorff, A. Greiner, R.B. Wehrspohn, K. Nielsch, J. Schilling, J. Choi, U. Gösele, *Science* **2002**, *296*, 1997.

[289] T.D. Blake, J. De Coninck, *Adv. Colloid Interface Sci.* **2002**, *96*, 21.

[290] P.G. de Gennes, A.M. Cazabat, *C.R. Acad. Sci.* **1990**, *310*, 1601.

[291] J. Bischof, D. Scherer, S. Herminghaus, P. Leiderer, *Phys. Rev. Lett.* **1996**, *77*, 1536.

[292] I. Strawbridge, P.F. James, *J. Non-Crystalline Solids* **1986**, *82*, 366.

[293] C.J. Lawrence, *Phys. Fluids* **1988**, *31*, 2786.

[294] A. Oron, S.H. Davis, S.G. Bankoff, *Rev. Mod. Phys.* **1997**, *69*, 931.

[295] M. Müller, L.G. MacDowell, P. Müller-Buschbaum, O. Wunnike, M. Stamm, *J. Chem. Phys.* **2001**, *115*, 9960.

[296] M.C. Fuerstenau, J.D. Miller, M.C. Kuhn, *Chemistry of Flotation*, Society of Mining Engineers, New York, **1985**.

[297] A.V. Nguyen, H.J. Schulze, *Colloidal Science of Flotation*, Marcel Dekker, New York, **2003**.

[298] T. Pfohl, S. Herminghaus, *Physik Journal* **2003**, *2*, 35.

[299] A. Manz, J.C.T. Eijkel, *Pure Appl. Chem*. **2001**, *10*, 1555.

[300] J. Khandurina, A. Guttman, *J. Chromatography A* **2002**, *943*, 159.

[301] H. Becker, C. Gärtner, *Electrophoresis* **2000**, *21*, 12.

[302] M. Wollmann, Dissertation, University of Siegen, **1996**.

[303] B. Berge, *C. R. Acad. Sci. Paris* **1993**, *317*, 157.

[304] J.A.M. Sondag-Huethorts, LG.J. Fokkink, *Langmuir* **1994**, *10*, 4380.

[305] E. Seyrat, R.A. Hayes, *J. Appl. Phys.* **2001**, *90*, 1383.

[306] M.W.J. Prins, W.J.J. Welters, Weekamp, J.W., *Science* **2001**, *291*, 277.

[307] G. Möller, M. Harke, H. Motschmann, D. Prescher, *Langmuir* **1998**, *14*, 4955.

[308] J. Lahann et al., *Science* **2003**, *299*, 371.

[309] M. Henzler, W. Göpel, *Oberflächenphysik des Festkörpers*, Teubner Studienbücher, Stuttgart, **1994**.

[310] K. Christmann, *Surface Physical Chemistry*, VCH Verlag, Weinheim, **1991**.

[311] H. Lüth, *Solid Surfaces, Interfaces and Thin Films*, $4^{th}$ ed., Springer, Berlin, **2001**.

[312] N. Harris, *Modern Vacuum Practice*, $2^{nd}$ ed., BOC Edwards, West Sussex, UK, **2001**.

[313] P.A. Redhead, *Vacuum* **1999** *53*, 137.

[314] M.P. Soriaga, *Prog. Surf. Sci.* **1992**, *39*, 325.

[315] R. Behrisch, K. Wittmaack (eds.), *Sputtering by Particle Bombardment, Vol. I,II,III*, Springer-Verlag, Berlin, **1981,1983,1991**.

[316] D. Axelrod, *J. Cell Biol.* **1981**, *89*, 141.

[317] A. Cho, *J. Vac. Sci. Tech.* **1971**, *8*, 31.

[318] A.Y. Cho (ed.), *Molecular Beam Epitaxy (Key papers in applied physics)*, AIP Press, New York, **1997**.

[319] M.A. Herman, H. Sitter, *Molecular Beam Epitaxy*, 2nd ed., Springer Series in Materials Science, Vol. 7, Springer, Berlin, **1996**.

[320] M. Ehsasi, K. Christmann, *Surf. Sci.* **1988**, *194*, 172.

[321] N. Camillone III, C.E.D. Chidsey, G. Liu, G. Scoles, *J. Chem. Phys.* **1993**, *98*, 4234.

[322] J. Morrison, J.J. Lander, *Surf. Sci.* **1966**, *5*, 163.

[323] R.J. Behm, K. Christmann, G. Ertl, M.A. Van Hove, *J. Chem. Phys.* **1980**, *73*, 2984.

[324] M. Yudin, B.D. Hughes, *Phys. Rev. B* **1994**, *49*, 5638.

[325] H.-J. Butt, R. Raiteri, in: *Surface Characterization Methods*, A.J. Milling (ed.), Marcel Dekker, New York, 1999, p. 1.

[326] R. Shuttleworth, *Proc. Phys. Soc. (London)* **1950**, *63A,* 444.

[327] P.W. Tasker, *Phil. Mag. A* **1979**, *39*, 119.

[328] G. Wulff, *Z. Krist.* **1901**, *34*, 449.

[329] K. Takayanagi, Y. Tanishiro, M. Takahashi, and S. Takahashi, *J. Vac. Sci. Techn. A* **1985**, *3*, 1502.

[330] A.L. Robinson, *Science* **1986**, *232,* 451.

[331] R. Lüthi, E. Meyer, M. Bammerlin, A. Baratoff, L. Howald, Ch. Gerber, H.-J. Güntherodt *Surf. Rev. Lett.* **1997**, *4*, 1025

[332] I. Stranski, *Z. Phys. Chem.* **1928**, *136*, 259.

[333] E.E. Gruber, W.W. Mullins, *J. Phys. Chem. Solids*, **1967** *28*, 875.

[334] P. Nozìeres, in: C. Godrèche (ed.) *Solids far from equilibrium*, Cambridge University Press, Cambridge, **1992**.

[335] W.K. Burton, N. Cabrera, F.C. Frank, *Philos. Trans. R. Soc. London, Ser. A* **1951** *243*, 299.

[336] J. Villain, D.R. Grempel, J. Lapujoulade, *J. Phys. F* **1985**, *15*, 809.

[337] M.S. Hoogeman, D.C. Schlößer, J.B. Sanders, L. Kuipers, J.W.M. Frenken, *Phys. Rev. B* **1996**, *53*, R13299.

[338] H. Jeong, E.D. Williams, *Surf. Sci. Rep.* **1999** *34*, 171.

[339] W.T. Read, *Dislocations in Crystals*, McGraw-Hill, New York, **1953**.

[340] W.T. Read, W. Shockley, *Phys. Rev.*, **1950**, *78*, 275.

[341] D.G. Brandon, B. Ralph, S. Ranganathan, M.S. Wald, *Acta Metall.*, **1964**, *12*, 813.

[342] V. Randle, *The Role of the Coincidence Site Lattice in Grain Boundary Engineering*, The Institute of Materials, London, **1996**.

[343] W. Bollmann, *Crystal Defects and Crystalline Interfaces*, Springer Verlag, Berlin, 1970.

[344] J.M. Howe, *Interfaces in Materials*, Wiley, New York, **1997**.

[345] J.P. Hirth, J. Lothe, *Theory of Dislocations*, $2^{nd}$ ed., Krieger Publishing, Malabar Florida, **1992**.

[346] D. Brune, R. Hellborg, H.J. Whitlow, O. Hunderi, *Surface Characterization*, Wiley-VCH, Weinheim, **1997**.

[347] J.T. Yates, *Experimental Innovations in Surface Science*, AIP Press, New York, **1997**.

[348] P.B. Garland, *Quart. Rev. Biophys.* **1996**, *29*, 91.

[349] S. Amelinckx, D. van Dyck, J. van Landuyt, G. van Tendeloo (eds.), *Electron Microscopy: Principles and Fundamentals*, Wiley-VCH, Weinheim, 1997.

[350] G.D. Danilatos, V.N.E. Robinson, *Scanning* **1979**, *2*, 72.

[351] M.K. Miller, A. Cerezo, M.G. Hetherington, G.D.W. Smith, *Atom Probe Field Ion Microscopy*, Clarendon Press, Oxford, 1996.

[352] G. Binnig, H. Rohrer, C. Gerber, E. Weibel, *Phys. Rev. Lett.* **1982**, *49,* 57.

[353] J. Tersoff, D.R. Hamann, *Phys. Rev. B* **1985**, *31,* 805.

[354] R. Garcia, R. Perez, *Surf. Sci. Rep.* **2002**, *47*, 197.

[355] D. Bonnell (ed.) Scanning Probe Microscopy and Spectroscopy, $2^{nd}$ ed., Wiley, New York, **2001**.

[356] R. Wiesendanger (ed.), *Scanning Probe Microscopy: Analytical Methods*, Springer, Berlin, **1998**.

[357] P. Broeckmann, M. Wilms, K. Wandelt, *Surf. REv. Lett* **1999**, *6*, 907.

[358] M. Deleu, M. Paquot, P. Jacques, P. Thonart, Y. Andriaensen, Y.F. Dufrene, *Biophys. J.* **1999**, *77*, 2304.

[359] H. Ibach, D.L. Mills, *Electron Energy Loss Spectroscopy and Surface Vibrations*, Academic Press New York, **1982**.

[360] A. Benninghoven, F.G. Rudenauer, H.W. Werner, *Secondary Ion Mass Spectrometry: Basic Concepts, Instrumental Aspects, Applications and Trends*, Wiley, New York, **1987**.

[361] A. Dabrowski, *Adv. Colloid Interface Sci.* **2001**, *93*, 135.

[362] M.A. Henderson, *Surf. Sci. Rep.* **2002**, *46*, 1.

[363] T.B. Grimley, in: *Interaction of Atoms and Molecules with Solid Surfaces*, V. Bortolani, N.H. March, M.P. Tosi (eds.), Plenum Press, New York, **1990**, p. 25.

[364] F.M. Devienne, in: *The Solid-Gas Interface Vol.2*, E.A. Flood (ed.), Marcel Dekker, New York, **1967**, p. 815.

[365] S. Brunauer, *The Adsorption of Gases and Vapors, Vol. 1*, Princeton University Press, Princeton, **1945**.

[366] H. Freundlich, *Kapillarchemie*, Akademische Verlagsgesellschaft, Leipzig, **1923**, Chapter 3.

[367] F. Rouquerol, J. Rouquerol, G.D. Gatta, C. Letoquart, *Thermochimica Acta* **1980**, *39*, 151.

[368] R.A. Pierotti, H.E. Thomas, in: *Surface and Colloid Science, Vol. 4*, E. Matijevic (ed.), Wiley-Interscience, New York, **1971**, p. 93.

[369] A.A. Isirikyan, A.V. Kiselev, *J. Phys. Chem.* **1961**, *65*, 601.

[370] I. Langmuir, *J. Am. Chem. Soc.* **1918**, *40,* 1361.

[371] D.E. Brown et al., *J. Phys. Chem.* **1996**, *100*, 4988.

[372] D.A. King, M.G. Wells, *Surf. Sci.* **1972**, *29*, 454.

[373] R.H. Fowler, *Proc. Camb. Phil. Soc.* **1935**, *31,* 260.

[374] A. Frumkin, *Z. physik. Chem.* **1925**, *116*, 466.

[375] R.H. Fowler, E.A. Guggenheim, *Statistical Thermodynamics*, Cambridge University Press, **1965**, p. 431.

[376] R.H. Fowler, *Proc. Camb. Phil. Soc.* **1936**, *32*, 144.

[377] Y. Larher, *J. Chem. Soc., Faraday Trans. I* **1974**, *70*, 320.

[378] S. Brunauer, P.H. Emmett, E. Teller, *J. Am. Chem. Soc*. **1938**, *60,* 309.

[379] H. Naono, M. Hakuman, *J. Colloid Interface Sci.* **1991**, *145*, 405.

[380] G. Halsey, H.S. Taylor, *J. Chem. Phys.* **1947**, *15*, 624.

[381] V.S. Mamleev, E.A. Bekturov, *Langmuir* **1996**, *12*, 441.

[382] W. Steele, *Appl. Surf. Sci.* **2002**, *196*, 3.

[383] G. Sauerbrey, Zeitschr. für Physik **1959**, *155*, 206.

[384] T. Nomura, M. IIjima, *Analytica Chimica Acta* **1981**, *131*, 97.

[385] C. Lu, A. Czanderna (eds.), *Applications of Piezoelectric Quartz Crystal Microbalances*, Elsevier, Amsterdam, **1984**.

[386] C.K. O'Sullivan, G.G. Guilbault, *Biosensors & Bioelectrocics* **1999**, *14*, 663.

[387] J.J.R. Stalgren, J. Eriksson, K. Boschkova, *J. Colloid Interface Sci.* **2002**, *253*, 190.

[388] H.G. Tompkins, *A User's Guide to Ellipsometry*, Academic Press, Boston, **1993**.

[389] J.W.O. Faul, U.G. Volkmann, K. Knorr, *Surf. Sci.* **1990**, *227*, 390.

[390] D.M. Young, A.D. Crowell, *Physical Adsorption of Gases*, Butterworths, London, **1962**.

[391] E.P. Barrett, L.G. Joyner, P.P. Halenda, *J. Am. Chem. Soc.* **1951**, *73*, 373.

[392] M.W. Cole, W.F. Saam, *Phys. Rev. Lett.* **1974**, *32*, 985.

[393] A. de Keizer, T. Michalski, G.H. Findenegg, *Pure & Appl. Chem.* **1991**, *63*, 1495.

[394] H. Liu, L. Zhang, N.A. Seaton, *J. Colloid Interface Sci.* **1993**, *156*, 285.

[395] M. Antonietti, *Curr. Opinion Colloid Interface Sci.* **2001**, *6*, 244.

[396] M.E. Davis, *Nature* **2002**, *417*, 813.

[397] C.T. Kresge, M.E. Leonowicz, W.J. Roth, J.C. Vartuli, J.S. Beck, *Nature*, **1992**, *359*, 710

[398] M. Kruk, M. Jaroniec, *Chem. Mater.* **2000**, *12*, 222.

[399] G.D. Parfitt, C.H. Rochester (ed.) *Adsorption from Solution at the Solid/Liquid Interface*, Academic Press, London, **1983**.

[400] G.J. Fleer, J.M.H.M. Scheutjens, in: *Coagulation and Flocculation*, B. Dobias (ed.), Marcel Dekker, New York, **1993**, p. 209.

[401] P. Somasundaran, L. Huang, *Adv. Colloid Interface Sci.*, **2000**, *88*, 179.

[402] R.G. Nuzzo, B.R. Zegarski, L.H. Dubois, *J. Am. Chem. Soc.* **1987**, *109*, 733.

[403] H.E. Kern, G.H. Findenegg, *J. Colloid Interface Sci.* **1980**, *75*, 346.

[404] J.P. Rabe, S. Buchholz *Science* **1991**, *253*, 424.

[405] H. Dautzenberg, W. Jaeger, J. Kötz, B. Philipp, Ch. Seidel, D. Stscherbina, *Polyelectrolytes – Formation, Characterization and Application*, Hanser Publishers, Munich, **1994**.

[406] B.D. Malhotra, A. Chaubey, *Sensors and Actuators B* **2003**, *91*, 117.

[407] W. Menz, J. Mohr, O. Paul, *Microsystem Technology*, Wiley-VCH, Weinheim, **2001**.

[408] H.J.P. Adler, K. Potje-Kamloth (eds.), *Quo vadis–coatings?* in: *Macromolecular Symposia* 187, Wiley-VCH, Weinheim, **2002**.

[409] S. PalDey, S.C. Deevi, *Materials Science and Engineering A* **2003**, *342*, 58.

[410] K.L. Choy, *Progress in Materials Science* **2003**, *48*, 57.

[411] A.R. Barron, *Advanced materials for optics and electronics* **1996**, *6(2)*, 101.

[412] S.K. Ray, C.K. Maiti, S.K. Lahiri, N.B. Chakrabarti, *Advanced materials for optics and electronics* **1996**, *6(2)*, 73.

[413] S.T. Lee, Z. Lin, X. Jiang, *Materials Science and Engineering* **1999**, *25*, 123.

[414] X. Jiang, K. Schiffmann, C.P. Klages, D. Wittorf, C.L. Jia, K. Urban, W. Jäger, *J. Appl. Phys.* **1998**, *83*, 2511.

[415] A. Ulman, *Chem. Rev.* **1996**, *96*, 1533.

[416] F. Schreiber, *Progr. Surf. Sci.* **2000**, *65*, 151.

[417] D.K. Schwartz *Ann. Rev. Phys. Chem.* **2001**, *52*, 107.

[418] R.G. Nuzzo, D.L. Allara, *J. Am. Chem. Soc.* **1983**, *105*, 4481.

[419] J. Xu, H.-L. Li, *J. Interface Colloid Sci.* **1995**, *176*, 138.

[420] C.D. Bain, J. Evall, G.M. Whitesides, *J. Am. Chem. Soc.* **1989**, *111*, 7155.

[421] M. Preuss, H.-J. Butt, *J. Colloid Interface Sci.* **1998**, *208*, 468.

[422] G. Nelles, H. Schönherr, M. Jaschke, H. Wolf, M. Schaub, J. Küther, W. Tremel, E. Bamberg, H. Ringsdorf, H.-J. Butt, *Langmuir* **1998**, *14*, 808.

[423] I. Haller, *J. Am. Chem. Soc.* **1978**, *100*, 8050.

[424] R. Jordan, in: M. Buchmeiser (ed.) *Polymeric materials in organic chemistry and catalysis*, Wiley-VCH, Weinheim, **2003**, in print.

[425] K.D. Lork, K.K. Unger, J.N. Kinkel *J. Chromatogr. A* **1986**, *352*, 199.

[426] J.D. Le Grange, J.L. Markham, *Langmuir* **1993**, *9*, 1749.

[427] M. Schönhoff, *Curr. Opinion Colloid Interface Science* **2003**, *8*, 86.

[428] G. Decher, Y. Lvov, J. Schmitt, *Thin Solid Films* **1994**, *244*, 772.

[429] K. Lowack, C.A. Helm *Macromolecules* **1998**, *31*, 823.

[430] B. Schwarz, M. Schönhoff, *Coll. Surf. A*, **2002**, *198*, 293.

[431] G. Decher, B. Lehr, K. Lowack, Y. Lvov, J. Schmitt, *Biosensors & Bioelectronics* **1994**, *9*, 677.

[432] E. Donath, G.B. Sukhorukov, F. Caruso, S.A. Davis, H. Möhwald, *Angewandte Chemie Intern. Ed.*, **1998**, *37*, 2202.

[433] A.A. Antipov, G.B. Sukhorukov, E. Donath, H. Möhwald *J. Phys. Chem. B*, **2001**, *105*, 2281.

[434] G. Ibarz, L. Dähne, E. Donath, H. Möhwald *Adv. Mater.* **2001**, *13(17)*, 1324.

[435] X. Shi, F. Caruso, *Langmuir* **2001**, *17(6)*, 2036.

[436] H.-G. Elias *An introduction to polymer science* Wiley-VCH, Weinheim, **1997**.

[437] O. Prucker, J. Rühe, *Macromolecules* **1998**, *31*, 592.

[438] O. Prucker, J. Rühe, *Macromolecules* **1998**, *31*, 602.

[439] C.J. Hawker *Acc. Chem. Res.*, **1997**, *30*, 373.

[440] T. Otsu, M. Yoshida *Makromol. Chem., Rapid Commun.* **1982**, *3*, 127.

[441] M. Kato, M. Kamigaito, M. Sawamoto, T. Higashimura, *Macromolecules* **1995**, *28*, 1721.

[442] V. Percec, B. Barboiu *Macromolecules* **1995**, *28*, 7970.

[443] J.-S. Wang, K. Matyjaszewski, *Macromolecules* **1995**, *28*, 7901.

[444] Y.J.T. Lii in: *ULSI Technology*, C.Y. Chang, S.M. Sze (eds.), McGraw-Hill, **1996**.

[445] W. Menz, P. Bley, *Mikrosystemtechnik für Ingenieure*, VCH, Weinheim, **1993**.

[446] C. Cardinaud, M.C. Peignon, P.Y. Tessier, *Appl. Surf. Sci.* **2000**, *164*, 72.

[447] C.M. Chan, *Polymer Surface Modification and Characterization*, Hanser/Gardner, Cincinnati, **1994**.

[448] F. Clement, B. Held, N. Soulem, H. Martinez, *Eur. Phys. J. AP* **2003**, *21*, 59.

[449] K.L. Prime, G.M. Whitesides, *J. Am. Chem. Soc.* **1993**, *115*, 10714.

[450] J.D. Gelorme, R.J. Cox, S.A.R. Gutierrez, *Photoresist composition and printed circuit boards and packages made therewith*, US patent 4,882,245, 21.Nov. **1989**

[451] N. Zhang, J. Xie, M. Guers, V.K. Varadan, *Smart Mater. Struct.* **2003**, *12*, 260.

[452] K.C. Ludema, *Friction, Wear, Lubrication: A Textbook in Tribology*, CRC Press, Boca Raton, **1996**.

[453] B. Bhushan, *Introduction to tribology*, John Wiley & Sons, New York, **2002**.

[454] J. Krim, *Scientific American,* Oct. **1996**, 48.

[455] B.N.J. Persson, *Sliding Friction: Physical Principles and Applications*, Springer, Berlin, **1998**.

[456] F.P. Bowden, D. Tabor, *Friction and Lubrication*, Methuen, London, **1954**.

[457] J. A. Greenwood, J. B. P. Williamson, *Proc. Roy. Soc. Lond.*, **1966**, *A295*, 300

[458] G.A. Tomlinson, *Phil. Mag. Series* **1929**, *7*, 905.

[459] F.P. Bowden, D. Tabor, *The Friction and Lubrication of Solids, Part 1 and 2*, Clarendon, Oxford, **1950** and **1964**.

[460] H. Czichos, *Tribology*, Elsevier, Amsterdam, **1978**.

[461] C.H. Scholz, *Nature* **1998**, *391*, 37.

[462] O. Reynolds, *Phil. Trans. Roy. Soc. London* **1876**, *116*, 155.

[463] H.L. Heathcote, *Proc. Inst. Automotive Engrs.* **1921**, *15*, 1569.

[464] N.V. Brilliantov, T. Pöschel, *Europhys. Lett.* **1998**, *42*, 511.

[465] D.G. Flom, A.M. Beuche, *J. Appl. Phys.* **1959**, *30*, 1725.

[466] T. Pöschel, T. Schwager, N. V. Brilliantov, *Eur. Phys. J. B*,**1999**, 10, 169

[467] S. Ecke, H.-J. Butt, *J. Colloid Interf. Sci.* **2001**, *244*, 432.

[468] C.M. Mate, G.M. McClelland, R. Erlandsson, S. Chiang, *Phys. Rev. Lett.* **1987** *59*, 1942.

[469] R.W. Carpick, M. Salmeron, *Chem. Rev.* **1997**, *97*, 1163.

[470] J.N. Israelachvili, P.M. McGuiggan, A. M. Homola, *Science* **1988**, *240*, 189.

[471] J. v. Alsten, S. Granick, *Phys. Rev. Lett.* **1988**, *61*, 2570.

[472] A. Dayo, W. Alnasrallah, J. Krim, *Phys. Rev. Lett.* **1998**, *80*, 1690

[473] J.S. Wettlaufer, J.G. Dash, *Scientific American*, **2000**, *282*, 50.

[474] E. Meyer, K. Dransfeld, R. Overney, T. Gyalog, *Nanoscience: Friction and Rheology on the Nanometer Scale*, World Scientific Publishing, Singapore, **1998**.

[475] G.V. Dedkov, *Phys. Stat. Sol.* **2000**, *3*, 179.

[476] L. Howald, R. Lüthi, E. Meyer, G. Gerth, H. Haefke, R. Overney, H.-J.Güntherodt, *J. Vac. Sci. Technol. B* **1994**, *12*, 2227.

[477] S. Fujisawa, E. Kishi, Y. Sugawara, S. Morita, *Phys. Rev. B* **1995**, *51*, 7849.

[478] H. Takano, M. Fujihira, *J. Vac. Sci. Technol. B* **1996**, *14*, 1272.

[479] R. Lüthi, E. Meyer, M. Bammerlin, L. Howald, H. Haefke, T. Lehmann, C. Loppacher, H.-J. Güntherodt, T. Gyalog, H. Thomas, *J. Vac. Sci. Technol. B* **1996**, *14*, 1280.

[480] T. Müller, T. Kässer, M. Labardi, M. Lux-Steiner, O. Marti, J. Mlynek, G. Krausch, *J. Vac. Sci. Technol. B* **1996**, *14*, 1296.

[481] H. Hölscher, U. D. Schwarz, O. Zwörner, R. Wiesendanger, *Phys. Rev. B* **1998**, *57*, 2477.

[482] K. Nakayama, B. Bou-Said, H. Ikeda, *Trans ASME J. Tribol.* **1997**, *119*, 764.

[483] F.J. Giessibl, M. Herz, J. Mannhart, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12006.

[484] O. Zwörner, H. Hölscher, U. D. Schwarz, R. Wiesendanger, *Appl. Phys. A* **1998**, *66*, 263.

[485] E. Gnecco, R. Bennewitz, T. Gyalog, Ch. Loppacher, M. Bammerlin, E. Meyer, H.-J. Güntherodt, *Phys. Rev. Lett.* **2000**, *84*, 1172.

[486] R. Bennewitz, T. Gyalog, M. Guggisberg, M. Bammerlin, E. Meyer, H.-J. Güntherodt *Phys. Rev. B* **1999**, *60*, R11302

[487] Y. Sang, M. Dubé, M. Grant, *Phys Rev. Lett.* **2001**, *87*, 174301.

[488] R.W. Carpick, N. Agraït, D. F. Ogletree, M. Salmeron, *J. Vac. Sci. Technol. B* **1996**, *14*, 1289.

[489] M.A. Lantz, S. J. O'Shea, M.E. Welland, K.L. Johnson, *Phys. Rev. B* **1997**, *55*, 10776.

[490] M. Enachescu, R.J.A. van den Oetelaar, R.W. Carpick, D.F. Ogletree, C.F.J. Flipse, M. Salmeron, *Phys.Rev. Lett.* **1998**, *81*, 1877.

[491] U.D. Schwarz, O. Zwörner, P. Köster, R. Wiesendanger, *Phys. Rev. B* **1997**, *56*, 6987 and 6997.

[492] P.M. McGuiggan, J. Zhang, S.M. Hsu, *Tribol. Lett.* **2001**, *10*, 217.

[493] D. Dowson, *History of Tribology*, Longman, New York, **1979**.

[494] O. Reynolds, *Phil. Trans. Roy. Soc. London*, **1886**, *177*, 157.

[495] B. Bhushan, J.N. Israelachvili, U. Landman, *Nature* **1995**, *374*, 607.

[496] Y. Hu, S. Granick, *Tribology Lett.* **1998**, *5*, 81.

[497] D.Y.C. Chan, R.G. Horn, *J. Chem. Phys.* **1985**, *83*, 5311.

[498] S. Granick, *Science* **1991**, *253*, 1374.

[499] J. Klein, E. Kumacheva, *J. Chem. Phys.*, **1998**, *108*, 6996.

[500] J. Klein, E. Kumacheva, *J. Chem. Phys.*, **1998**, *108*, 7010.

[501] M. Sun, C. Ebner, *Phys. Rev. Lett.* **1992**, *69*, 3491.

[502] P.A. Thompson, and S.M. Troian, *Nature* **1997**, 389, 360.

[503] M.J. Stevens, M. Mondello, G.S. Grest, S.T. Cui, H.D. Cochran, P.T. Cummings, *J. Chem. Phys.* **1997**, *106*, 7303.

[504] O.I. Vinogradova, *Int. J. Miner. Process.*, **1999**, *56*, 31.

[505] J. Baudry, E. Charlaix, A. Tonck, D. Mazuyer, *Langmuir*, **2001**, *17*, 5232.

[506] R.G. Horn, O.I. Vinogradova, M.E. Mackay, N. Phan-Thien, *J. Chem. Phys.*, **2000**, *112*, 6424.

[507] X. Zhu, S. Granick, *Phys. Rev. Lett.* **2001**, *87*, 096105.

[508] V.S.J. Craig, C. Neto, D.R.M. Williams, *Phys. Rev. Lett.* **2001**, *87*, 054504.

[509] G. Sun, E. Bonaccurso, V. Franz, H.-J. Butt, *J. Chem. Phys.* **2002**, *117*, 10311.

[510] T. Mang, W. Dresel (eds.), *Lubricants and Lubrication*, Wiley-VCH Weinheim, **2001**.

[511] M. Rasberger, in: *Chemistry and Technology of Lubricants*, R. M. Mortimer, S. T. Orszulik (eds.), Blackie Academic and Professional, **1992**, p. 83.

[512] J.F. Archard, *J. Appl. Phys.*, **1953**, *24*, 981.

[513] S.H. Meng, PhD research, Mechanical Engineering Department, University of Michigan, Ann Arbor, **1994**.

[514] P.H. Shipway, I. M. Hutchings, *Wear*, **1996** *193*, 105.

[515] A.V. Levy, Solid Particle Erosion and Erosion-Corrosion of Materials, ASM International, **1995**.

[516] C. Kajdas, *Wear* **1985**, *101*, 1.

[517] R.G. Laughlin, *The Aqueous Phase Behavior of Surfactants*, Academic Press, London, **1996**.

[518] W. von Rybinski, K. Hill, *Angew. Chem. Int. Ed.* **1998**, *37*, 1328.

[519] P.D.T. Huibers, V.S. Lobanov, A.R. Katritzky, D.O. Shah, M. Karelson, *J. Colloid Interface Sci.*, **1997**, *187*, 113.

[520] P.D.T. Huibers, V.S. Lobanov, A.R. Katritzky, D.O. Shah, M. Karelson, *Langmuir* **1996**, *12*, 1462.

[521] R. Zana, *Adv. Colloid Interface Sci.* **2002**, *97*, 205.

[522] N. Hadjichristidis, S. Pispas, G.A. Floudas, *Block Copolymers*, Wiley-Interscience, Hoboken, **2003**.

[523] P. Stilbs, H. Walderhaug, B. Lindman, *J. Phys. Chem.* **1983**, *87*, 4762.

[524] W.M. Gelbart, A. Ben-Shaul, D. Roux, *Micelles, Membranes, Microemulsions and Monolayers*, Springer, New York, **1994**.

[525] E.A.G. Aniansson, S.N. Wall, M. Almgren, H. Hoffmann, I. Kielmann, W. Ulbricht, R. Zana, J. Lang, C. Tondre, *J. Phys. Chem.* **1976**, *80*, 905.

[526] F. Krafft, H. Wiglow *Berichte d. Dtsch. chem. Gesellschaft* **1895**, *28*, 2573.

[527] U. Herrmann, M. Kahlweit, *Ber. Bunsenges. Phys. Chem.* **1973**, *77*, 1119.

[528] P. Mukerjee, *Adv. Colloid Interface Sci.* **1967**, *1*, 241.

[529] R. Zana, *Langmuir* **1996**, *12*, 1208.

[530] D.F. Evans, *Langmuir* **1988**, *4*, 3.

[531] C. Tanford, *The Hydrophobic Effect*, Wiley, New York, 1980.

[532] J.B. Jeong, S.R. Yang, J.D. Kim, *Langmuir* **2002**, *18*, 8749.

[533] J.N. Israelachvili, D.J. Mitchell, B.W. Ninham, *J. Chem. Soc. Faraday Trans II* **1976**, *72*, 1525.

[534] H.-F. Eicke, *Topics in Current Chemistry* **1980**, *87*, 85.

[535] M. Born, *Zeitschr. f. Physik* **1920**, *1*, 45.

[536] S. Friberg, *Food Emulsions*, Marcel Dekker, New York, **1976**.

[537] P. Barham, *The Science of Cooking*, Springer, Berlin, **2001**.

[538] P. Becher, *Emulsions: Theory and Practice*, Reinhold Publishing, New York, **1965**.

[539] R.E. Meredith, C.W. Tobias, *J. Electrochem. Soc.* **1961**, *108*, 286.

[540] G. Gompper, M. Schick, *Self-assembling Amphiphilic Systems*, Academic Press, London, **1994**.

[541] A. Einstein, *Annalen der Physik* **1906**, *19*, 289; **1911**, *34*, 591.

[542] W.D. Bancroft, *J. Phys. Chem.* **1913**, *17*, 501; **1915**, *19*, 275.

[543] G.H.A. Clowes, *J. Phys. Chem.* **1916**, *20*, 407.

[544] W.C. Griffin, *J. Soc. Cosmet. Chem.* **1949**, *1*, 311; **1954**, *5*, 249.

[545] K. Shinoda, S. Friberg, *Emulsions and Solubilization*, Wiley-Interscience, New York, **1986**.

[546] H.T. Davis, *Colloids & Surfaces A* **1994**, *91*, 9.
[547] M. von Smoluchowski, *Phys. Z.* **1916**, *17*, 557; 585.
[548] D.E. Tambe, M.M. Sharma, *Adv. Colloid Interface Sci.* **1994**, *52*, 1.
[549] P. Aussillous, D. Quéré, *Nature* **2001**, *411*, 924.
[550] W.D. Harkins, *The Physical Chemistry of Surface Films*, Reinhold Publishing, New York, **1957**, p. 90.
[551] A. Kabalnov, H. Wennerström, *Langmuir* **1996**, *12*, 276.
[552] T.P. Hoar, J.H. Schulman, *Nature* **1943**, *152*, 102.
[553] P.A. Winsor, *Trans. Faraday Soc.* **1948**, *44*, 376.
[554] J.E. Bowcott, J.H. Schulman, *Zeitschr. f. Elektrochemie* **1955**, *59*, 283.
[555] J.H. Schulman, W. Stoeckenius, L.M. Prince, *J. Phys. Chem.* **1959**, *63*, 1677.
[556] J. Sjöblom, R. Lindberg, S.E. Friberg, *Adv. Colloid Interface Sci.* **1996**, *95*, 125.
[557] W. Helfrich, *Z. Naturforschung* **1973**, *28c*, 693.
[558] R. Lipowsky, *Encyclodedia of Appl. Phys.* **1998**, *23*, 199.
[559] A. Kabalov, H. Wennerström, *Langmuir* **1996**, *12*, 276.
[560] K. Shinoda, H. Sagitani, *J. Colloid Interface Sci.* **1978**, *64*, 68.
[561] D. Vollmer, J. Vollmer, R. Strey, *Phys. Rev. E* **1996**, *54*, 3028.
[562] M. Kahlweit, R. Strey, *Angew. Chem.* **1985**, *97*, 655.
[563] R. Strey, *Ber. Bunsenges. Phys. Chem.* **1993**, *97*, 742.
[564] R.J. Pugh, *Adv. Colloid Interface Sci.* **1996**, *64*, 67.
[565] C. Isenberg, *The Science of Soap Films and Soap Bubbles*, Dover Publications, New York, **1992**.
[566] D. Weaire, S. Hutzler, *The Physics of Foams*, Clarendon Press, Oxford, **1999**.
[567] J. Banhart, *Progress in Material Science* **2001**, *46*, 559.
[568] J.K. Cochran, *Current Opinion in Solid State & Materials Science*, **1998**, *3*, 474.
[569] E. Manegold, *Schaum*, Chemie und Technik Verlagsgesellschaft, Heidelberg, **1953**.
[570] V. Bergeron, *J. Phys. Condens. Matter* **1999**, *11*, R215.
[571] K.J. Mysels, M.N. Jones, *Discuss. Faraday Soc.* **1966**, *42*, 42.
[572] D. Exerova, A. Scheludko, *Chim. Phys.* **1971**, *24*, 47.
[573] V. Bergeron, C.J. Radke, *Langmuir* **1992**, *8*, 3020.
[574] D. Langevin, A.A. Sonin, *Adv. Colloid Interface Sci.* **1994**, *51*, 1.

[575] A. Pockels, *Nature*. **1891**, *43*, 437.
[576] I. Langmuir, *J. Am. Chem. Soc.* **1917**, *39*, 1848.
[577] J.W. McBain, T.F. Ford, D.A. Wilson, *Kolloid-Z.* **1937**, *78*, 1.
[578] J.W. McBain, J.R. Vinograd, D.A. Wilson, *J. Am. Chem. Soc.* **1940**, *62*, 244.
[579] H. Möhwald, C. Böhm, A. Dietrich, S. Kirstein, *Liquid Crystals*. **1993**, *14*, 265.
[580] H.M. McConnell, *Ann. Rev. Phys. Chem.* **1991**, *42*, 171.
[581] C.M. Knobler, R.C. Desai, *Ann. Rev. Phys. Chem.* **1992**, *43*, 207.
[582] W.D. Harkins, T.F.Young, E. Boyd, *J. Chem. Phys.* **1940**, *8*, 954.
[583] W.D.Harkins, L.E. Copeland, *J. Chem. Phys.* **1942**, *10*, 272.

[584] H. Möhwald, *Ann. Rev. Phys. Chem.* **1990**, *41*, 441.

[585] C.A. Helm, P. Tippmann-Krayer, H. Möhwald, J. Als-Nielsen, K. Kjaer, *Biophys. J.* **1991**, *60*, 1457.

[586] K. Graf, *Dissertation* (in German), Mainz, **1997**.

[587] G.L. Gaines, *Insoluble Monolayers at Liquid-Gas Interfaces*, Interscience Publ., New York, **1966**.

[588] H. Möhwald in: *Phospholipid Handbook*, G. Cevc (ed.), Marcel Dekker Inc., New York, **1993**, p. 579.

[589] M. Lösche, H. Möhwald, *Rev. Sci. Instrum.* **1984**, *55*, 1968.

[590] M. Lösche, E. Sackmann, H. Möhwald, *Ber. Bunsenges. Phys. Chem.* **1983**, *87*, 848.

[591] R.M. Weis, H.M. McConnell, *Nature*. **1984**, *310*, 47.

[592] H.M. McConnell, *Ann. Rev. Phys. Chem.* **1991**, *42*, 171.

[593] D. Hönig, D. Möbius, *J. Phys. Chem.* **1991**, *95*, 4590.

[594] S. Hénon, J. Meunier, *Rev. Sci. Instrum.* **1991**, *62*, 936.

[595] C. Lheveder, S. Hénon, R. Mercier, G. Tissot, P. Fournet, J. Meunier, *Rev. Sci. Instrum.* **1998**, *69*, 1446.

[596] M. Harke, R. Teppner, O. Schulz, H. Motschmann, H. Orendi, *Rev. Sci. Instrum.* **1997**, *68*, 3130.

[597] D. Vollhardt, *Adv. Colloid Interface Sci.* **1996**, *64*, 143.

[598] S. Rivière, S. Hénon, J. Meunier, D.K. Schwartz, M.W. Tsao, C.M. Knobler, *J. Chem. Phys.* **1994**, *101*, 10045.

[599] J. Als-Nielsen, H. Möhwald, in: *Handbook on Synchroton Radiation, Vol. 4*, S. Ebashi, M. Koch and E. Rubenstein (eds.), Elsevier Science Publishers, **1991**, p. 1.

[600] C.A. Helm, P. Tippmann-Krayer, H. Möhwald, J. Als-Nielsen, K. Kjaer, *Biophys. J.* **1991**, *60*, 1457.

[601] K. Kjaer, *Physica B* **1994**, *198*, 100.

[602] A. Braslau, M. Deutsch, P.S. Pershan, A.H. Weiss, J. Als-Nielsen, J. Bohr, *Phys. Rev. Lett.* **1985**, *54*, 114.

[603] D.K. Schwartz, M. L. Schlossmann, E. H. Kawamoto, G. J. Kellog,P. S. Pershan, B. M. Ocko, *Phys. Rev. A* **1990**, *41*, 5687.

[604] A. Braslau, M. Deutsch, P.S. Pershan, A.H. Weiss, J. Als-Nielsen, J. Bohr, *Phys. Rev. Lett.* **1985**, *54*, 114.

[605] K. Graf, H. Baltes, H. Ahrens, C.A. Helm, C.A. Husted, *Biophys. J.* **2002**, *82*, 896.

[606] B. Ohler, I. Revenko, C. Husted, *J. Struct. Biol.* **2001**, *133*, 1.

[607] C.A. Helm, H. Möhwald, K. Kjaer, J. Als-Nielsen, *Biophys. J.* **1987**, *52*, 381.

[608] Z. Li, M.W. Zhao, J. Quinn, M.H. Rafailovich, J. Sokolov, R.B. Lennox, A. Eisenberg, X.Z. Wu, M.W. Kim, S.K. Sinha, M.Tolan, *Langmuir* **1995**, *11*, 4785.

[609] S.W. Barton, B.N. Thomas, E.B. Flom, S.A. Rice, B. Lin, J.B. Peng, J.B. Ketterson, P. Dutta, *J. Chem. Phys.* **1988**, *89*, 2257.

[610] R.M. Kenn, C. Böhm, A.M. Bibo, I.R. Peterson, H. Möhwald, J. Als-Nielsen, K. Kjaer, *J Phys. Chem.* **1991**, *95*, 2092.

[611] F. Paltauf, H. Hauser, M.C. Phillips, *Biochim. Biophys. Acta* **1971**, *249*, 539.

[612] V. Vogel, D. Möbius, *Thin Solid Films* **1988**, *159*, 73.
[613] A. Geil, *Dissertation* (in German), Mainz, **1995**.
[614] O. N. Oliveira, C. Bonardi, *Langmuir* **1997**, *13*, 5920.
[615] R.J. Demchak, T. Fort, *J. Colloid Interface Sci.* **1974**, *46*, 191.
[616] E.F. Porter, *J. Am. Chem. Soc.* **1937**, *59*, 1883.
[617] H.G. Yamins, W.A. Zisman, *J. Chem. Phys.* **1933**, *1*, 656.
[618] K.D. Wantke, K. Lunkenheimer, C. Hempt, *J. Colloid Interface Sci.* **1993**, *159*, 28.
[619] H. Bercegol, J. Meunier, *Nature* **1992**, *356*, 226.
[620] I. Langmuir, *Trans. Faraday Soc.* **1920**, *15*, 62.
[621] K.B. Blodgett, *J. Am. Chem. Soc.* **1935**, *57*, 1007.
[622] G. Roberts, *Langmuir-Blodgett-Films*, Plenum Press, New York, **1990**.
[623] A. Ulman, *An Introduction to Ultrathin Organic Films from Langmuir-Blodgett to Self-Assembly*, Academic Press, San Diego, **1991**.
[624] M.C. Petty, *Langmuir-Blodgett-Films: An Introduction*, Cambridge University Press, Cambridge, **1996**.
[625] D.K. Schwartz, *Surf. Sci. Rep.* **1997**, *27*, 241.
[626] J.G. Petrov, H. Kuhn, D. Möbius, *J. Coll. Int. Sci.* **1980**, *73*, 66.
[627] K. Spratte, H. Riegler, *Makromol. Chem., Macromol. Symp.* **1991**, *46*, 113.
[628] H. Riegler, K. Spratte, *Thin Solid Films*, **1992**, *210/211*, 9.
[629] K. Graf, H. Riegler, *Coll. Surf. A* **1998**, *131*, 215.
[630] W. Kurz, D.J. Fisher, *Fundamentals of solidification*, $4^{th}$ ed., Trans Tech Publications, **2001**.
[631] J.D. Swalen, D.L. Allara, J.D. Andrade, E.A. Chandross, S. Garoff, J. Israelachvili, T.J. McCarthy, R. Murray, R.F. Pease, J.F. Rabolt, K.J. Wynne, H. Yu, *Langmuir* **1987**, *3*, 932.
[632] I.R. Peterson, in: *Molecular Electronics: Properties, Dynamics, and Applications*, G. Mahler, V. May, M. Schreiber (eds.), Marcel Dekker, New York, **1996**.
[633] A.A. Freitas, F.H. Quina, F.A. Carroll, *J. Phys. Chem. B* **1997**, *3*, 7488.

# Index


Abrasion 243
Accommodation coefficient 179
Activator 140
Adhesion 114, 129, 132
Adhesion force 112, 115
Adhesive 133
Adhesive contact 243
Adsorbate 177
Adsorpt 177
Adsorption
    energy, 202
    isotherm, 178, 179, 181, 185, 195
    polymer, 181
    time, 178, 179
Adsorptive 177
Advancing contact angle 128
Aerosol 2
AerosolOT 247
AES 172, 173
AFM 97, 166, 230, 256
AgCl 63
Aggregation number 251, 252
AgI 63, 64
$Al_2O_3$ 65, 66
Alcohol 105
Alkyl glycoside 247
Alkylbenzene sulfonate 247
Alkylcarboxylate 247
Alkyldimethylpropanesultaine 247
Alkylethylene glycol 247
Alkylglucoside 247
Alkylsulfate 247, 254, 255
Alkyltrimethylammonium bromide 247
Alkyltrimethylammonium chloride 247
Amontons' law of friction 224
Amphiphile 242, 246
Amphoteric surfactant 247
Anionic surfactant 246
Annealing 151
Antifoamer 242
Antioxidant 242
Antiwear additive 242
AOT 247
Aquaplaning 236
Aquatic birds 125
Archard's wear equation 241
Asperity 225
Atomic beam diffraction 170
Atomic force microscope 97, 106, 109, 116, 166, 230
Auger electron spectroscopy 172
AW additive 242
Azobenzene 143

Bancroft's rule 264
Barrel etching 218, 219
Bending elastic modulus 270
Bending rigidity 270
Benzene 185
BET adsorption isotherm 181, 189, 192, 195, 205
Bilayer 257
Biological membrane 258
Black films 276
Blood plasma 45
Bolaform surfactant 250
Boltzmann equation 43
Boron nitride CVD film 209
Bragg condition 321, 323
Bravais lattice 147
Brewster angle microscopy 287
Bridging force 110
Butter 263

Calorimeter 183



*Physics and Chemistry of Interfaces.* Hans-Jürgen Butt, Karlheinz Graf, Michael Kappl


Copyright © 2003 Wiley-VCH Verlag & Co. KGaA


ISBN: 3-527-40413-9

Capacity 50, 53, 56, 59, 70, 71
Capillary 141, 143
Capillary condensation 17, 18, 20, 200
Capillary constant 12, 125
Capillary force 18–20
Capillary number 134
Capillary penetration technique 126
Capillary rise 122, 123
Capillary rise method 126
Capillary wave 289
Captive bubble 126
Cassie equation 131
Cationic surfactant 246
Cavitation 244
Centrifuge 114, 115
Cetyltrimethylammonium bromide 107
Champagne 23
Chemical vapor deposition 207–209
Chemisorption 178, 201, 202
Chronoamperometry 71
Clay 66
Cleavage 151
Cloud point 252
CMC 38, 140, 250–255
Coagulation 265
Coalescence 266, 267
Coefficient of friction 224
Coherent interface 160
Coincident site lattice model 161
Cold welding 226
Collector 140
Colloid 1
Colloidal probe technique 97, 114, 128, 229, 235
Composite material 116
Condensation coefficient 187
Conductivity 260
Conductometric titration 70
Confined liquid 104, 105, 240
Confocal microscopy 162
Constant charge 100
Constant potential 100
Contact angle 12, 118
Contact angle hysteresis 128
Continuous phase 259
Core region 131
Corrosion inhibitor 242
Cosurfactant 270
Coulomb force 80, 98
Coulomb friction 223
Coulomb's law of friction 226
Creaming 266, 267
Critical micellar concentration 38
Critical radius 21
Critical temperature 178
CSL model 161
CTAB 107, 246, 247, 256
Curvature 11
CVD 207
Cyclic voltametry 72
Cylindrical micelle 256

DDAB 247
De-inking of paper 139
Debye interaction 83
Debye length 45, 49, 68
Debye–Hückel theory 42, 76
Defects 157
Demulsification 267
Demulsifier 242
Depressant 140
Derjaguin approximation 93, 95, 101
Desorption energy 202, 205
Detergency 133, 140
Detergent 246
Dewetting 139
Dialkyldimethylammonium bromide 247
Diamond CVD film 208
Diblock copolymer 250
Diffuse layer 42
Dimeric surfactant 250
Dip coating 138
Dipole moment 81
Dipole–dipole interaction 82, 85
Disjoining pressure 95, 96, 100, 101
Dislocation 159
    edge, 159
    screw, 159
Dislufide 209
Dispersant 242
Dispersing agent 259
Dispersion 1, 2
Dispersion interaction 83
DLVO force 103, 106
DLVO theory 103, 104
DMT model 113, 114
Double-layer force 98, 101, 103, 131, 265
Double-wedge technique 151

DPPC 172, 285, 286
Drop-weight method 13, 24
Dropping mercury electrode 61
Dry friction 223
Du-Noüy tensiometer 14
Dynamic friction 223
Dynamic viscosity 240

Earthquake 227
EDX 172, 173
Electric double layer 42, 49, 50, 54, 57, 77, 98, 264
Electro-osmosis 73–75, 142
Electrocapillarity 58, 60, 78
Electrocapillary curve 60, 62
Electrochemical cell 64, 78
Electrochemical potential 78
Electrophoresis 76, 142
Electrophoretic retardation 77
Electrowetting 142
Ellipsometer 197
Ellipsometry 136, 196
Emulsifier 263, 265
Emulsion 2, 123, 242, 259, 260
Energy dispersive X-ray analysis 172
Energy of adsorption
  differential, 184
  integral, 182
Enthalpy 33
Enthalpy of adsorption 183
  differential, 184
  integral, 182
Entropy 33
Entropy of adsorption
  differential, 184
  integral, 182
EP additive 242
Erosion 243
ESCA 171
Etching 217
  barrel, 218, 219
  dry, 217
  ion beam, 217
  plasma, 218, 219
  reactive ion, 218, 219
  reactive ion beam, 218
  sputter, 218
Evaporation 152
Ewald construction 325–327
EXAFS 172
Extended X-ray absorption fine structure 172
External phase 259
Extreme pressure additive 242

Fatty acid 247
$Fe_3O_4$ 66
Fermi energy 173
FFG 189
FFM 230
Fibre 125
Film 137
Film balance 281
Film pressure 281
Flocculation 265, 266
Flotation 89, 124, 133, 138
Fluid lubrication 236
Fluorescence microscopy 286
Foam 2
Forced wetting 133
Form factor 328, 329
Fretting wear 243
Freundlich adsorption isotherm 180, 193
Friction 223
Friction force 115
Friction force microscope 230
Froth 139
Frothing agent 140
Frumkin–Fowler–Guggenheim isotherm 189, 190

Galactocerebroside 289
Galvani potential 58, 63, 77, 78
Gecko 87
Gemini surfactant 250
Germanium 68
Gibbs 26
Gibbs adsorption isotherm 36, 38, 60, 264
Gibbs dividing plane 26, 28, 29, 32, 181
Gibbs energy of adsorption 183, 188
Gibbs energy of micellization 253, 254
Gibbs free energy 32, 54
Gibbs interfacial energy 34
Gibbs monolayer 280, 282
Gibbs surface energy 34
Gouy–Chapman layer 42, 43
Gouy–Chapman theory 42, 49, 51, 71
Grafting-from polymerization 216
Grahame equation 49, 50, 64, 76

Grain boundary 160
tilt, 160
twist, 161
Grazing incidence X-ray diffraction 170
Guggenheim 26

Hagen–Poiseuille's law 141
Hamaker constant 86, 88, 89, 91, 132, 238
Heat 35
Heat of adsorption 179, 182, 183, 189
HEED 170
Helfrich 269
Helium 92
Helmholtz free energy 82
Helmholtz interfacial energy 33
Helmholtz layer 42, 43, 53
Henry adsorption isotherm 180, 195
Hertz model 111, 112, 116, 225
Heterogeneous surface 129, 130, 192
Heterophase interface 160
High energy electron diffraction 170
High internal phase emulsion 260
HIPE 260
HLB scale 264
Homogeneous nucleation 20
Homophase interface 160
HOPG 151
HREELS 173
Hydration force 106, 264
Hydraulic radius 199
Hydrophobic effect 254
Hydrophobic interaction 107

Ice 232
Ideal freely jointed segments chain 108
Ideal interface 26, 27, 38
Iep 72
Impedance spectroscopy 72
Inclined plane tribometer 230
Incoherent interface 160
Ink 108
Ink bottle pore 199–201
Inner Helmholtz plane 53
Inner phase 259
Inner phase volume fraction 260
Interface 1
Interfacial enthalpy 33
Interfacial excess 27, 58
Interfacial tension 32
Interference microscopy 162
Internal energy 29
Internal interfacial energy 32
Internal surface energy 32, 35, 36
Intersegment force 110
Inverse gas chromatography 157
Inverted micelle 257
Ion beam etching 217
Ion etching 218
Ion plating 206
Ionizing electrode method 292
Isoelectric point 72

JKR model 112–116, 229, 230

Keesom interaction 82
Kelvin equation 15–17, 21
Kelvin probe 291
Kinematic viscosity 240
Kinetic friction 223
Krafft point 252
Krafft temperature 252

Lab-on-chip 141
Lamellar phase 257
Langmuir adsorption isotherm 181, 185, 186, 189
Langmuir constant 181, 186, 189
Langmuir equation 186, 187, 189
Langmuir trough 281
Langmuir–Blodgett transfer 172, 293
Laplace equation 8
Laplace pressure 8, 15
Lateral force microscope 230
Laue condition 324, 326
Launderometer 140
LB film 293, 294
LEED 169, 170
LFM 230
Lifshitz theory 87
Line tension 121, 287
Lipid 257
Lipid bilayer 92, 258
Lippman equation 59
London interaction 83
Lotus effect 130
Low pressure chemical vapor deposition 207
Low-energy electron diffraction 169
LPCVD 207

Lubricant 133, 240
Lubrication 223, 236–238
    boundary, 236, 238
    elasto-hydrodynamic, 237
    hydrodynamic, 236, 237
    mixed, 238
    thin film, 240

Macroemulsion 259–262, 265, 266
Macropore 199
Marangoni effect 40, 265
Maximum-bubble-pressure method 13
MBE 153
Mean free path 207
Membrane protein 259
Membrane resistance, specific 258
Meniscus force 18, 20, 24
Mercury 60, 62, 64
Mercury porosimeter 198
Mesophase 283
Mesopore 199
Mica 67, 96, 116
Micelle 140, 242, 251–254, 256
Micellization 253
Microcontact 225, 229, 230
Microemulsion 259
Microfluidic 141, 142
Micropore 199
Microsphere tensiometry 127
Milk 260
Miller indices 146, 324
Mixed lubrication 236
$MnO_2$ 66
Modulus of Gaussian curvature 270
Molecular beam epitaxy 153
Molecular-kinetic theory 136
Monolayer 181, 280, 283, 285–289, 291–294
    insoluble, 280

Nanocontact 235
Nanotribology 232
Navier–Stokes equation 73
Nearest neighbor broken bond model 155
Nernst equation 63
Nonionic surfactant 247
Nucleation rate 21
Null ellipsometer 197

O-lattice theory 161
Oil 240
Oil-in-water emulsion 260, 261, 264
Optical microscopy 162
Osmotic stress method 97
Outer Helmholtz plane 53
Outer phase 259
Oxidation 241
Oxide 65, 69

Packing ratio 255
Paint 133, 137
Particle 18, 19, 87, 123, 127, 139
PECVD 207
Pendant bubble 13
Pendant drop 12
Phase inversion temperature 265, 271
Phospholipid 280, 281
Physical vapour deposition 152
Physisorption 178, 202
Pin-on-disk tribometer 230
Pinning 129
PIT 265, 271
Plasma enhanced chemical vapor deposition 207
Plasma etching 218, 219
PLAWM trough 281, 282
Plunging-tape experiment 134
Point of zero charge 62, 66, 69, 72
Poisson equation 43
Poisson–Boltzmann equation 44, 47, 48, 100
Poisson–Boltzmann theory 43, 50, 100
Polanyi theory of adsorption 193, 195
Polarizability 82
Polarizable electrode 64
Polyallylamine 213
Polyelectrolyte 212, 215
Polyethyleneimine 213
Polylysine 213
Polymer 65, 107–110, 212, 215, 263
Polymer brush 109
Polymer film 138, 139
Polymeric surfactant 250
Polystyrenesulfonate 213
Porous material 198
Porous solid 2, 18
Potential determining ion 63, 65
Potential theory 193, 195
Potentiometric titration 69, 70
Pour point depressant 242

Powder 114, 126, 182, 198, 263
Precursor film 135, 136
Primary energy minimum 103
Principal curvature 11
Propanol 105
Protein 65
Pulp 139
PVD 152
Pzc 62

QCM 196, 231
Quartz crystal microbalance 153, 196, 231

Radius of curvature 8, 9
Radius of gyration 108
Reactive ion etching 218, 219
Receding contact angle 128
Reciprocal lattice 324
RED 170
Relative adsorption 28
Retardation 84
Reversible electrode 64, 71
Reynolds number 141, 237
RHEED 170
Ring tensiometer 14
Rolling friction 228
Roughness 128–130

Saddle-splay modulus 270
SAM 212, 213
Scanning electron microscope 163, 164
    environmental, 163
Scanning force microscope 97, 165
Scanning tunneling microscope 165
Scattering vector 322
SDS 37, 39, 142, 246, 247, 250, 251, 253, 254, 256
Secondary energy minimum 103, 104, 265
Secondary ion mass spectrometry 174
Sedimentation potential 77
SEM 163
Semicoherent interface 160
Semiconductor 68
Sessile bubble 126
Sessile bubble method 13
Sessile drop 12, 126
Sessile drop method 12
SFA 96, 231
SFM 165
Shear rate 237
Shear thinning 237, 238
Shuttleworth equation 154
Silane 210
Silanization 210, 211
Silicon 68
Silver chloride 63
Silver iodide 63
SIMS 174
    dynamic, 174
    static, 175
Sintering 145
$SiO_2$ 65, 66, 87, 116
Size of a polymer 108
Slippage 240
Soap 246
Sol 2
Solid friction 223
Solubility 253
Solvation force 105
Specific surface area 182
Spin coating 138
Spinning drop method 296
SPM 164
Spontaneous curvature 269
Spontaneous spreading 133
Spreading coefficient 122, 295
Sputter etching 218
Sputtering 151, 152
Step 159
Steric force 107, 109, 264
Stern layer 52
Stick-slip friction 226, 227
    atomic, 233
Sticking probability 187
STM 165
Stokes law 76
Streaming potential 75
Stribeck diagram 238, 239
Structure factor 328
Substrate structure 146
Super-water-repellent surface 130
Superlattice 149
Superposition principle 81
Surface charge 43, 49, 66, 69, 76, 99
Surface charge density 59
Surface elasticity 292
Surface energy 113
    solid, 153, 155, 156

Surface enthalpy 33, 36
Surface entropy 33–36
Surface excess 29
Surface fatigue 243
Surface forces apparatus 96, 97, 105, 109, 115, 231
Surface intensive parameter 153
Surface melting 151
Surface potential 43, 49, 63, 65, 76, 77, 99
monolayer, 290–292
Surface reconstruction 147
Surface relaxation 147
Surface strain 153
Surface stress 153, 156, 157
Surface tension 5–7, 12, 31, 32, 36
solid, 153, 156
Surface viscosity 292
Surfactant 37, 140, 246, 262, 264
Surfactant number 256
Surfactant parameter 255, 257
Suspending power 140
Suspension 2

Tapping mode 167
TDS 202, 205
Tears of wine 39
Teflon 92
TEM 163
Temperature programmed desorption 202
Tenside 246
Terrace-ledge-kink model 157
Terraces 157
Thermal decomposition 241
Thermal desorption spectroscopy 202
Theta solvent 108
Theta temperature 108
Thin film 89
Thin film balance 276
Thiol 131, 143, 150, 202, 209
Three-phase contact line 118, 121, 128, 129, 133
$TiO_2$ 65, 66
TOF-SIMS 175
TPD 202
Transmission electron microscope 163
Tree 123
Tribochemical reactions 243
Tribology 223
Tribometer 230
Trimeric surfactant 250

UHV 145
Ultraviolet photon spectroscopy 78
UPS 173
UV photoemission spectroscopy 173

Vacuum 206, 207
Van der Waals equation of state 84
Van der Waals force 89, 91, 93, 96, 103, 105
Van der Waals interaction 83–86, 131, 138, 155, 156, 178, 193–195, 265, 267, 286
Vapor pressure 15, 17
Vibrating electrode method 291
Vicinal surface 157, 158
Viscosity 237, 240, 260
Viscosity index 241
Viscosity modifier 242
Volatility 241
Volta potential 77

Water purification 139
Water repellent ability 125
Water-in-oil emulsion 260, 261, 264
Wave vector 321
Wear 223, 241, 243
Wenzel equation 130
Wetting line 118, 121
Wilhelmy plate method 14, 24, 126, 282
Wood's notation 150
Work function 78
Wulff construction 154

X-ray diffraction 287–289
X-ray photoemission spectroscopy 171
X-ray reflection 287–289
XPS 171, 173

Yield stress 225
Young's equation 118–122, 131, 295
Young–Laplace equation 8, 10–12, 31

Zeta potential 52, 72, 74, 75
Zwitterionic surfactant 247