Theory of soft and biomatter

Ulrich Schwarz

June 8, 2009

Contents

1 Introduction **3**
- 1.1 What is soft matter ? . . . 3
- 1.2 Physical scales of biomaterials . . . 5
- 1.3 Further reading . . . 10

2 Molecular and colloidal interactions **13**
- 2.1 Covalent bonding . . . 13
- 2.2 Coulomb interaction . . . 14
  - 2.2.1 Point charges . . . 14
  - 2.2.2 Distributed charges . . . 16
  - 2.2.3 Counterions in solution . . . 19
  - 2.2.4 One charged surface . . . 21
  - 2.2.5 Two charged surfaces . . . 23
- 2.3 Dipolar and van der Waals interactions . . . 26
- 2.4 Hydrophilic and hydrophobic interactions . . . 31
- 2.5 Colloidal dispersions . . . 33

3 Simple and complex fluids **36**
- 3.1 Ideal gas . . . 36
- 3.2 Virial expansion for real gas . . . 38
  - 3.2.1 Hard spheres . . . 40
  - 3.2.2 Attractive energy . . . 41
  - 3.2.3 Square well potential . . . 41
- 3.3 Fluid-fluid phase transition . . . 42
- 3.4 Fluid-solid phase transition . . . 43
- 3.5 Simple fluids . . . 45
- 3.6 Multiple component fluids . . . 46
- 3.7 Liquid crystals . . . 47

4 Interfaces **49**
4.1 ϕ^4-model 49
4.2 Differential geometry of surfaces 54
4.3 Capillary waves 59
4.4 Rayleigh-Plateau instability 60
4.5 Tension-related phenomena 62

5 Membranes **67**
5.1 Bending energy 67
5.2 Thermal fluctuations 69
5.3 Helfrich-interaction 71
5.4 Vesicles 72

6 Elasticity **76**
6.1 Strain tensor 76
6.2 Stress tensor 78
6.3 Elastic moduli 79
6.4 Some simple cases 82
6.4.1 Pure (hydrostatic) compression 82
6.4.2 Pure shear 83
6.4.3 Uniaxial stretch 83
6.5 Beams and thin shells 84
6.6 Capsules 84
6.7 Contact mechanics 86

7 Hydrodynamics **89**
7.1 Fundamentals 89
7.2 Simple examples 92
7.2.1 Shear (Couette) flow 92
7.2.2 Plane Poiseuille flow 93
7.2.3 Cylindrical Poiseuille flow 94
7.3 Stokes flow 94
7.3.1 Hydrodynamic lubrication 94
7.3.2 Viscous adhesion 96
7.3.3 Stokes drag for a sphere 97

Chapter 1

Introduction

1.1 What is soft matter ?

This question is easy to answer: soft matter distorts easily in response to external force. However what exactly is meant with *easy* ? Which physical scales are we talking about ? In order to address this question, it is instructive to consider the simple stretching experiment depicted in Fig. 1.1. A bulk sample of elastic material of resting length L_0 is stretched to the extension $L_0 + \Delta L$ by a force couple F applied at two of its faces over a cross-sectional area A. Therefore the deformation is caused by the stress $\sigma = F/A$ and can be characterized by the strain $\epsilon = \Delta L/L_0$. For small deformations, one usually finds a linear relationship between stress and strain

$$\sigma = E\epsilon \tag{1.1}$$

where E is the elastic modulus (*Young modulus*) representing the material properties. This implies that the elastic slab behaves like a Hookean spring, with a linear force-displacement relation $F = K\Delta L$, where the spring constant is given by $K = EA/L_0$. Because strain is dimensionless, the dimension of the Young modulus is the same as for stress, that is $[\mathrm{E}] = [\sigma] = \mathrm{Pa} = \mathrm{N/m^2} = \mathrm{J/m^3}$. The Young modulus therefore has to scale as $E = U/a^3$, with U being a typical energy and a being a typical length scale of the system under consideration. Thus soft matter has to be characterized either by a small energy scale or by a large length scale.

In practice, most soft matter systems show both features. With regard to energy, soft condensed matter is held together by non-covalent interactions

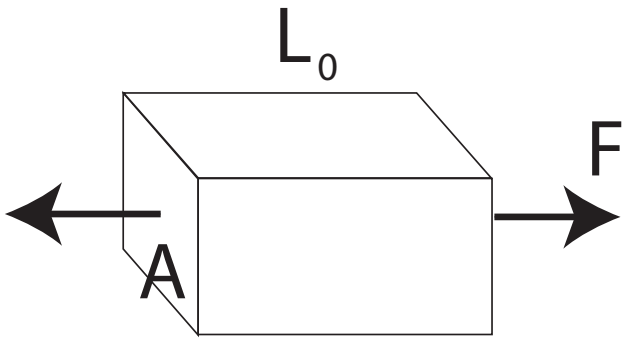


Figure 1.1: A slab of elastic material with crossectional area A and resting length L_0 is stretched by ΔL by applying a force couple F. The mechanical response of the material is described by the elastic modulus $E = FL_0/A\Delta L$.

like van der Waals forces or hydrogen bonds. Then the typical energy scale is thermal energy k_BT, where k_B is the Boltzmann constant $k_B = 1.38 \times 10^{-23}$ J/K and T is ambient temperature. A convenient choice for scaling arguments is $T = 300$ K, so that k_BT becomes $4.1 \times 10^{-21}J = 4.1$ pN nm. In a physical chemistry context, it is more common to refer to macroscopic quantities, so we use the Avogadro number $N_A = 6.022 \times 10^{23}$ to compute $k_BTN_A = RT = 2.5$ kJ/mol $= 0.6$ kcal/mol, where $R = k_BN_A$ is the molar gas constant.

With regard to length, soft matter usually corresponds to supramolecular aggregates, thus typical length scales are much larger than atomic ones. For example, a colloidal crystal is characterized by thermal energy $U = k_BT$ and a supramolecular length $a = 10$ nm, resulting in an elastic modulus $E = kPa$. In contrast, atomic crystals, which are traditionally studied in solid-state physics, are characterized by $U = eV = 40$ kT and an atomic length $a = 0.1$ nm, resulting in $E = GPa$. More extreme values are the Pa-range for very soft gelatine and the TPa-range for carbon nanotubes, while the usual houseware rubber in the MPa-range represents the middle ground.

Due to the small energy scale, soft matter systems are intrinsically dynamic and thermal noise is sufficient to induce structural changes. This outstanding property of soft matter systems allows them to self-organize and to self-heal. On the one side, this endows soft matters systems which a sort of robustness which usually is not present in traditional condensed matter systems (e.g. usually an atomic crystal after fracture has no chance to grow together again). On the other hand, it also makes them very fragile, because

small perturbations can easily perturb the system.

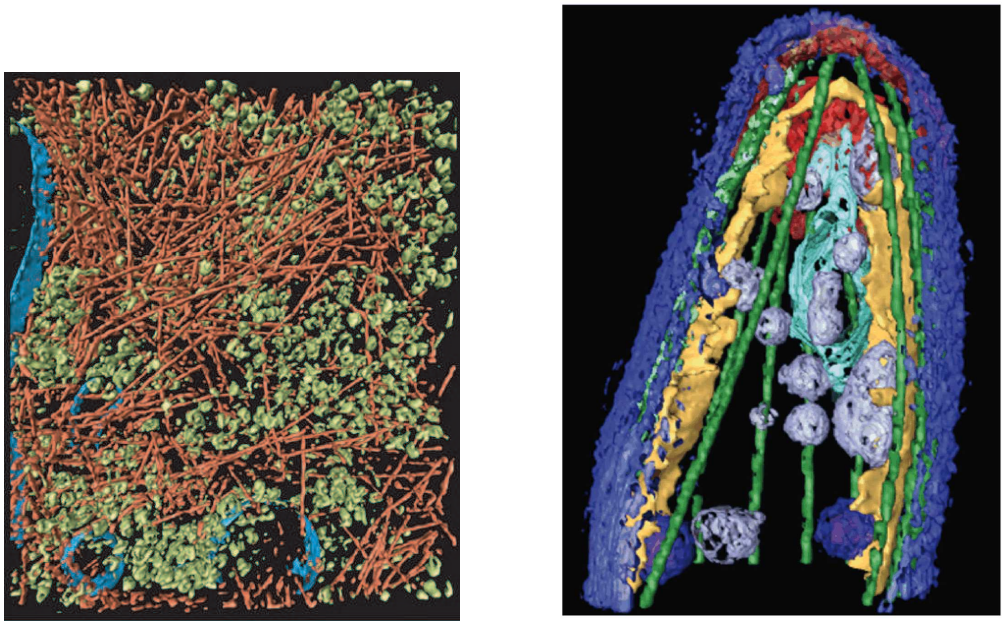
Soft matter systems are very important due to their technological relevance. In fact, many such systems are known to us from everyday life, including all kinds of bath room items (tooth paste, shampoo, cosmetics, etc), food (salad dressing, yoghurt, pudding, ketchup, etc), paint, the liquid crystals in our computer displays, motor oil and beer foam. In general, the term *soft matter* includes material systems like colloidal suspensions, liquid crystals, fluid-fluid interfaces, fluid membranes and polymers, all of which are treated in detail in the following.

Throughout this text, special emphasis will be put on the fact that all biological systems are made of soft matter, including the cells and tissues in our body. The structural properties of cells are mainly determined by lipid membranes and the filamentous proteins of the cytoskeleton. Tissue is a composite material comprising cells and an extracellular protein network (*extracellular matrix*). All fundamental processes of life rely on the fact that biological systems are made from soft matter, because its dynamic character allows the systems to adapt quickly to changes in their environment.

However, it is important to also note that biological systems differ in essential aspects from inanimate soft matter systems, for three related reasons. First, biological systems are the result of several billion years of evolution and are under strict genetic control. This implies that many genetically encoded programmes can be triggered in the cell which leads to very specific behaviour that cannot be predicted on physical grounds. Second, generic interactions between the different components are suppressed in favor of specific interactions, for example van der Waals-attraction between cell surfaces is suppressed in favor of the lock-and-key reactions between surface receptors and their ligands. And third, biological systems are active in the sense that with the help of metabolic reservoirs, active processes take place which would not occur spontaneously. In contrast to traditional condensed matter systems, soft matter allows biological systems to tightly couple structural and regulatory processes.

1.2 Physical scales of biomaterials

The marvels of nature are one important motivation to study the physics of soft matter. As a background to the main part of this text, we now give some more details regarding the physical scales in biological systems.



(a) (b)

Figure 1.2: State-of-the-art electron tomography of cellular structures. (a) This image shows in red the actin cytoskeleton of a migrating amoeba. Ribosomes and membranes are shown in green and blue, respectively. (b) This image shows the upper part of a malaria parasite at the stage at which it is transmitted from the mosquito into the host skin. The microtubules in green are the main stabilizing element for cell shape.

The quantum of life is the cell (here we assume that autonomous replication is one criterion for life, so a virus is not considered to represent life). As Rudolf Virchow realized in 1858, every cell originates from the division of another cell. Cells are small, but not inconceivably small. For human cells, the typical size is 10 μm. Therefore 10^3 cells fit into 1 cm and $(10^3)^3 = 10^9$ cells fit into a finger tip. That is roughly as many as rice corns fit into a middle-sized room. Interestingly, the size of cells (10 μm) behave to macroscopic scales (1 cm) as the size of large biomolecules (10 nm) to the cellular scale (10 μm). Therefore 10^9 of these large biomolecules could make up one cells, each with around 10^5 atoms. This estimate then results in $10^5 10^9 10^9 = 10^{23}$ atoms on the macroscopic scale, exactly as requested by Avogadro's number.

Since the naked eye resolution is 200 μm, we cannot see cells. In order

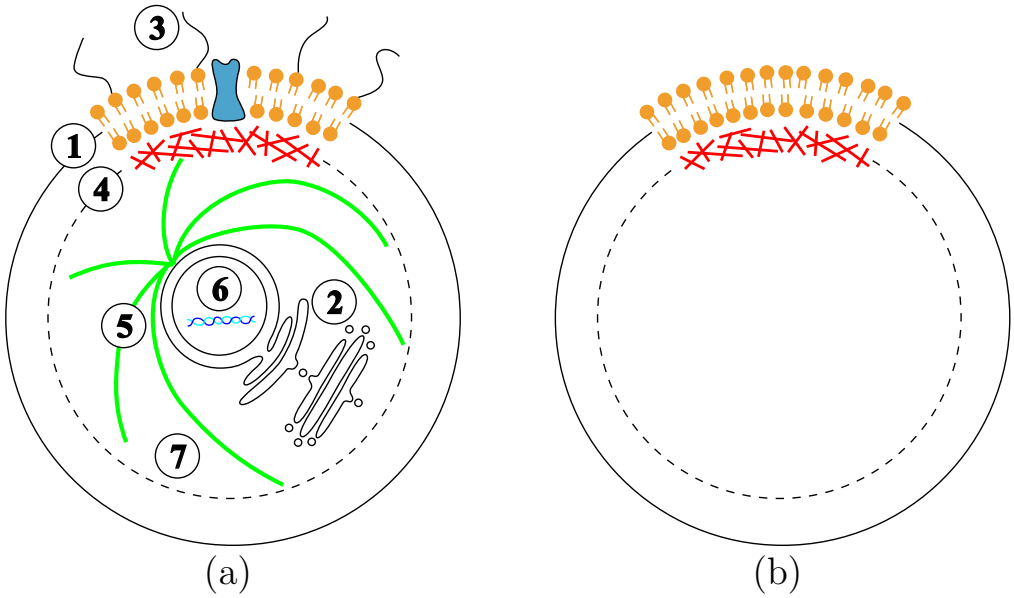


Figure 1.3: (a) Schematic drawing of a higher animal cell. Important cellular structures include (1) the plasma membrane, (2) other membrane structures like the two nuclear membranes, the endoplasmic reticulum, the Golgi apparatus and vesicles, (3) the glycocalix, (4) the actin cortex, (5) the microtubule system, (6) DNA and (7) the cytoplasm. (b) In a biomimetic approach to cell shape and mechanics, one can reduce the structural complexity of the cell by considering only the effects of plasma membrane and actin cortex. Such a reduced system in fact corresponds quite well to the situation with the red blood cell, which therefore is the most important model system to study cell shape and mechanics.

to do so, we need a microscope, so in 1665, Robert Hooke was the first to see cells. In fact optical resolution is 200 nm and by far sufficient to observe cells. In order to look at subcellular structures in detail, one needs an electron microscope, with a resolution better than nm. This is sufficient to see lipid bilayers, globular proteins and protein filaments, which all have sizes of a few nm. Nowadays the most impressive visualizations of cellular structures result from electron tomography. Fig. 1.2 shows the visualization of the lamellipodium of a migrating amoeba.

As already mentioned, the typical energy scale for soft matter is $k_BT = 4.1$ pN nm. With a typical length scale of nm, this leads to a typical force scale of pN for proteins. As explained above, a typical elastic scale for

supramolecular aggregates is kPa. In fact this is also the order of magnitude measured with the atomic force microscope for the stiffness of animal cells like the fibroblasts from connective tissue (the detailed values depend on cell type and the current state of the cell; bacteria and viruses are much stiffer). For single protein domains, the relevant length scale (e.g. the distance between van der Waals bonds) is rather smaller and the elastic scale reaches the GPa-range.

Due to their common origin, all cells are assembled from similar subcellular components. The sophisticated structures of cells from higher animals like humans are of particular interest for biomedical reasons. In order to characterize their physical properties experimentally, these properties have to be defined in an operative way. During recent years, many new experimental techniques have been developed which now allow to study the physical properties of living cells, including fluorescent probes, atomic force microscopy, laser optical tweezers, (optical) cell stretchers, colloidal bead microrheology, elastic substrates and microfluidic chips. Alternatively, biomaterial can be reconstituted and probed in the test tube, which allows to compare with macroscopic measurement techniques. In the following, we list a few of the subcellular structures of interest (a schematic drawing is provided in Fig. 1.3a) and the physical scales which have been found in experiments. The detailed meaning of these physical quantities will be explained later in this text:

(1) plasma membrane: a lipid bilayer which serves as a space partitioner and permeability barrier; due to its fluidity, transmembrane proteins can move freely in this two-dimensional sheet; a good model are vesicles; overall bilayer thickness 4 nm, hydrocarbon layer thickness 2 nm, with a dielectric constant of 2, in contrast to 80 for the surrounding water; bilayer bending rigidity 20 k_BT, rupture tension 10^{-3} J/m^2, stretching modulus 0.2 J/m^2

(2) other membrane structures: other structures formed by lipid bilayers are the two nuclear membranes, the endoplasmic reticulum, the Golgi apparatus, transport vesicles, endosomes, and mitochondria

(3) glycocalix: a polyelectrolyte brush surrounding the cell, controls adhesion and ligand binding; strongly depends on cell type, typical thickness 50 nm, most surface receptors are buried in it

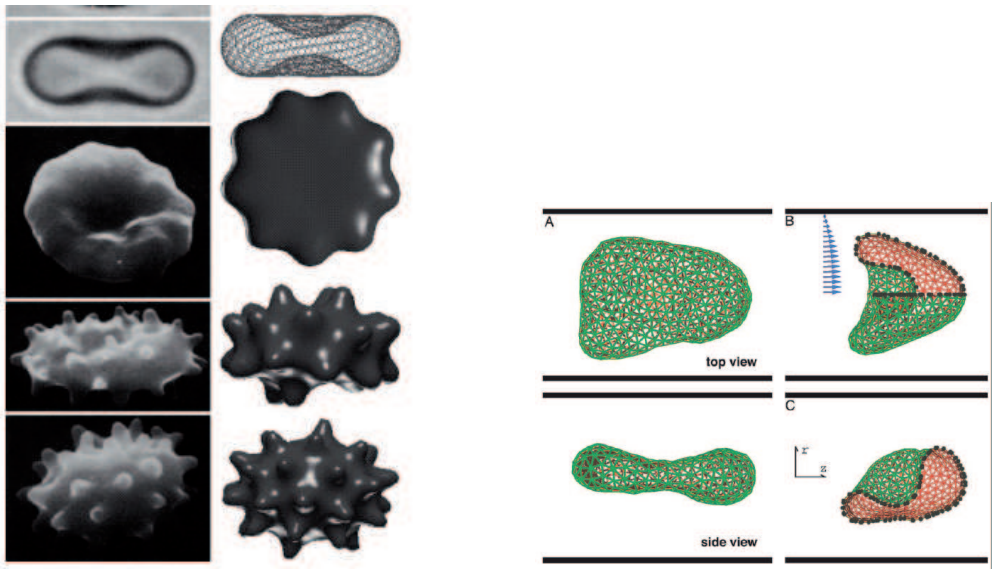
(4) actin cortex: a thin layer of actin polymer network located below the plasma membrane; filamentous actin is a semiflexible polymer dynamically crosslinked by a variety of additional proteins; together with myosin II molecular motors, it is also the essential component of muscle; a good model for the actin cortex are polymeric capsules; thickness 100 nm, elastic modulus kPa; the composite sandwich of plasma membrane and actin cortex determines much of cell mechanics, which therefore has to be described by thin shell elasticity; in experiments, it was found that the cell envelope is under an effective tension of the order of nN/μm

(5) microtubuli: a system of stiff polymers radiating from a common center, the microtubule organizing center close to the nucleus; microtubules provide the tracks along which molecular motors like dynein and kinesin transport their cargo, and contribute to overall cell mechanics; actin filaments and microtubules together with intermediate filaments (not shown here) form the cytoskeleton, that is the system of polymer networks which gives structural integrity to eukaryotic cells

(6) DNA: a flexible polymer whose basepair sequence carries the genetic information; human DNA has a length of 2×3 giga basepairs (Gbp) $\times$ 0.34 nm = 2 m, which is compactified into the μm-sized nucleus despite the high charge of $2e^-$ per basepair, a feat which is accomplished mainly by multivalent counterions like spermidine and with the help of histones, large proteins onto which the DNA is wrapped

(7) cytoplasm: the dense solution of ions and molecules filling the interior of the cell; the viscosity for the movement of small particles is only modestly larger than the one of water, but the macroscopic viscosity (which is relevant e.g. when deforming the whole cell) is 10^4 Pa s (similar to honey, 10^7 the one of water); salt concentration 100 mM (like sea water), which leads to a Debye screening length of only 1 nm

The cellular system depicted in Fig. 1.3a is too complex as to be approached head-on with physical methods. In order to investigates the physical aspects of cells, it is helpful to reduce the cellular complexity and to start with appropriately reduced models. In Fig. 1.3b, the reduced version of Fig. 1.3a is shown which already reproduces many important aspects of cell shape and mechanics, namely a polymer capsule wrapped by a lipid bilayer.



(a) (b)

Figure 1.4: Physical modelling of red blood cells. (a) The equilibrium shapes of red blood cells as a function of area difference between the two monolayers of the lipid bilayer. Electron micrographs on the left fit nicely to theoretical results on the right. (b) Typical shapes of red blood cells in hydrodynamic shear flow as simulated with multiple particle collision dynamics.

In fact such a system exists in nature, namely the red blood cell, which has lost most of its intracellular inventory when converted into the transporter vehicle as which it is used in the body. In Fig. 1.4 we show results from the physical modelling of red blood cell shape both in equilibrium and in shear flow. These results demonstrate that the physical principles for cell shape and mechanics are indeed amendable to theoretical analysis. One of the purposes of this text is to introduce the concepts required to deal with these issues.

1.3 Further reading

For the theory of soft matter physics, there are several excellent books available, including

- SA Safran, Statistical thermodynamics of surfaces, interfaces, and mem-

branes, Addison-Wesley, Reading 1994

- PM Chaikin and TC Lubensky, Principles of condensed matter physics, Cambridge University Press, Cambridge 1995
- KA Dill and S Bromberg, Molecular driving forces : statistical thermodynamics in chemistry and biology, Garland Science 2003

The following proceedings from the winter schools at Research Center Juelich provide an up-do-date view on soft matter research:

- JKG Dhont, G Gompper and D Richter, eds, Soft Matter: Complex Materials on Mesoscopic Scales, Proceedings winter school Research Center Juelich 2002
- G Gompper, UB Kaupp, JKG Dhont, D Richter and RG Winkler, eds, Physics meets Biology: From Soft Matter to Cell Biology, Proceedings winter school Research Center Juelich 2004

The following books combine experimental and theoretical issues of soft matter systems:

- JN Israelachvili, Intermolecular and surface forces, 2nd edition, Academic Press 1992
- DF Evans and H Wennerström, The colloidal domain: where physics, chemistry, and biology meet, 2nd edition, Wiley 1998
- RAL Jones, Soft condensed matter, Oxford University Press 2002
- T Witten and P Pincus, Structured fluids : polymers, colloids, surfactants, Oxford 2004

For elasticity theory, the standard textbook is

- LD Landau and EM Lifschitz, Elasticity Theory (VII), Pergamon Press, Oxford, 1986

For polymer theory, the standard textbooks are

- M Doi and SF Edwards, The theory of polymer dynamics, Clarendon Press, Oxford, 1986

- M Rubinstein and RH Colby, Polymer Physics, Oxford University Press 2003

For membranes, vesicles and capsules, excellent reviews are provided in

- R Lipowsky and E Sackmann, Eds., Structure and Dynamics of Membranes, Elsevier, Amsterdam 1995
- U Seifert, Configurations of fluid membranes and vesicles, Advances in Physics 46: 13-137, 1997.
- C Pozrikidis, ed., Modeling and Simulation of Capsules and Biological Cells, Chapman and Hall / CRC Press, 2003

Finally there is an increasing number of books on biological physics:

- J Howard, Mechanics of motor proteins and the cytoskeleton, Sunderland 2001
- D Boal, Mechanics of the cell, Cambridge University Press 2002
- P Nelson, Biological physics: energy, information, life, Freeman 2004
- R Philips, J Kondev and J Theriot, Physical Biology of the Cell, Taylor and Francis 2008

Chapter 2

Molecular and colloidal interactions

In this chapter we will learn that soft matter is characterized by interactions that are weak, attractive and short-ranged. In particular, we will see that cohesion in soft matter systems usually arises from the electrostatic interaction. A detailed analysis shows that this one kind of interaction leads to a long list of seemingly uncorrelated phenomena. The fact that today we can understand them all from a common source shows how well developed our understanding of soft matter systems has become over time. We will start with the Coulomb interaction between charged particles and then address the issue of distributed charge. This includes mobile counterions in solution, which leads to screening and short-ranged interactions. Then we turn to dipolar interactions, including the ubiquitous van der Waals interaction. We will discuss hydrophilic, hydrophobic and depletion interactions, which are important to understand both colloidal and biological systems. Finally we discuss the nature of biological interactions, which roughly speaking correspond to a complicated mixture of soft matter interactions to achieve very specific interactions.

2.1 Covalent bonding

Covalent or chemical bonding arises from the electron cloud shared between different atoms and its strength has to be calculated from quantum mechanics. In general, covalent bonding is short-ranged, directional and strong. For

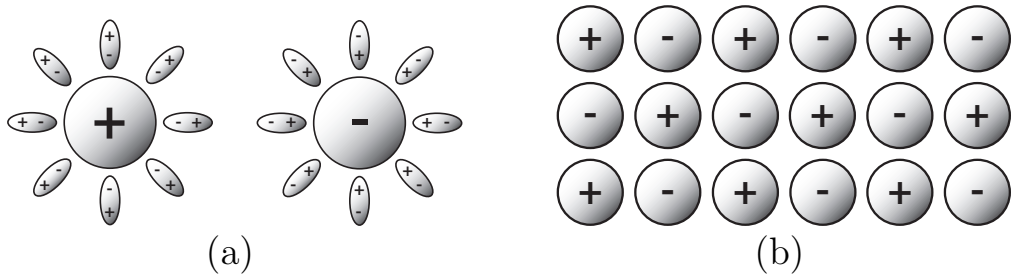


Figure 2.1: (a) The attraction between two oppositely charged particles in solution is weakened if the solvent molecules are polarizable. This effect is described by the dielectric constant ϵ (2 or oil and 80 for water). (b) A charge-neutral assembly of positive and negative charges can be stable, for example the sodium chloride crystal (Na^+Cl^-) schematically shown here.

single, double and triple carbon-carbon bonds, one finds energies of 140, 240 and 330 k_BT, respectively. Covalent bonding in condensed matter systems depends on the details of the chemical environment. For example, the C-H bond in H-CHO and H-CN has energies of 144 and 200 k_BT, respectively. Metallic bonding is similar to covalent bonding and also has an energy scale much larger than thermal energy (band structure in solids is usually calculated in eV = 40 k_BT).

2.2 Coulomb interaction

2.2.1 Point charges

The pairwise Coulomb or ionic interaction energy reads

$$U = \frac{q_1 q_2}{4\pi\epsilon_0\epsilon}\frac{1}{r} \tag{2.1}$$

where the q_i are the two charges and r is distance. $\epsilon_0 = 8.85\times10^{-12}$ C^2/Jm is the permittivity of vacuum and ϵ is the dielectric constant of the surrounding medium, with approximate values of 1, 2 and 80 for air, hydrocarbon and water. A large dielectric constant means that the Coulomb interaction is weakened because the medium surrounding the charges is polarized, compare Fig. 2.1a. It is important to note that the concept of the dielectric constant ϵ is not simply an assumption, but rather it is a non-trivial result obtained by

spatial averaging over regions of $\approx (10\ nm)^3$. Usually the dielectric constant decreases with increasing temperature (i.e., the Coulomb interaction becomes stronger), because thermal motion decreases the degree of polarization. The $1/r$-dependance in Eq. (2.1) implies that the Coulomb-interaction is long-ranged, like the gravitational interactions which accounts for the structure of the universe.

In general, ionic interactions are as strong as chemical bonding. For example, for one sodium (Na^+) and one chloride (Cl^-) ion one has an equilibrium distance at room temperature and in vacuum of $r = 0.28$ nm. The charges are simply unit charges, $q_1 = -q_2 = e = 1.6 \times 10^{-19}$ C. Using Eq. (2.1), one finds $U = -200\ k_BT$, that is the same order of magnitude as for covalent bonding. If one now considers the sodium chloride (Na^+Cl^-) crystal, the Coulomb interaction from Eq. (2.1) has to be summed over all ions in the crystal, compare Fig. 2.1b. Surprisingly, the result for the energy per ion differs only by a factor of 1.747, the so-called *Madelung constant*. The main effect at work here is that positive and negative charges more or less cancel each other, a phenomenon called *screening*. The theoretical result also shows that charge-neutral arrangements can be stable, meaning that they can provide a favorable cohesive energy. In macroscopic units, it corresponds to a cohesive energy of -206 kcal/mol. Experimentally one finds -181 kcal/mol, which is surprisingly close to the theoretical result. The dielectric constant ϵ does not enter here because the vacuum is taken to be the reference state. However it becomes important if one wants to explain why a salt crystal can dissolve in water: because here ϵ is large, going into solution does not cost much energy and thus there is always a sizable probability for ions to be in solution, where they gain entropy.

Soft matter and biological systems are often characterized by the presence of surfaces or interfaces. In many cases, their effect can be described by introducing so-called *image charges*. Image charges are placed outside the sample volume in such a way that the boundary conditions are just right. The classical example for the interaction of a point charge with a surface is the case when this surface is conducting, compare Fig. 2.2a. Then the electric field has to be normal to the surface everywhere, otherwise charges in the conductor would move. In the case of a plane conducting surface, one simply has to assume an image charge of *opposite* sign, $q' = -q$, opposing the original charge q. Therefore the surface is effectively attractive. In soft matter physics, however, conducting surfaces are rare and a much more relevant case is an interface with a jump in the dielectric constant. An important

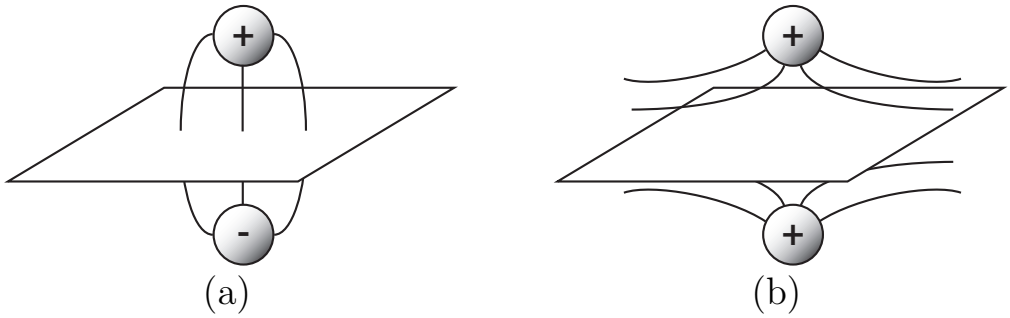


Figure 2.2: (a) A charge close to a conductor induces an image charge of opposite sign and therefore is attracted by the surface. (b) Close to an interface to low dielectric constant like oil, an image charge of equal sign is induced and repulsion results.

example is the water-oil interface with $\epsilon_w = 80$ and $\epsilon_o = 2$. The situation of a point charge in the water phase is solved by two image charges, namely a charge $q' = q(\epsilon_w - \epsilon_o)/(\epsilon_w + \epsilon_o) \approx q$ at the opposing position and a charge $q'' = 2q\epsilon_o/(\epsilon_w + \epsilon_o) \approx 0$ at the same position. Therefore one essentially has an image charge of *equal* sign and repulsion results, compare Fig. 2.2b. This is another example of the general tendency that charges prefer the medium of high dielectric constant.

2.2.2 Distributed charges

In soft matter and biological systems, charges are often spatially distributed. In order to study how interaction laws vary with geometry in the case of the Coloumb interaction, we start with a short review of electrostatics. From the interaction potential Eq. (2.1), the force on a particle follows as

$$\mathbf{F} = -\nabla U = \frac{q_1 q_2}{4\pi\epsilon_0\epsilon}\frac{\mathbf{r}}{r^3} = q_2\mathbf{E} \tag{2.2}$$

where $\mathbf{E}$, the force on a test particle with positive unit charge, is the electrostatic field. Experimentally, one finds that the electrostatic field arising from different charges is a superposition of the electrostatic fields which result from the single charges:

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0\epsilon}\int d\mathbf{r}'\rho(\mathbf{r}')\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3} = -\nabla\phi(\mathbf{r}) \tag{2.3}$$

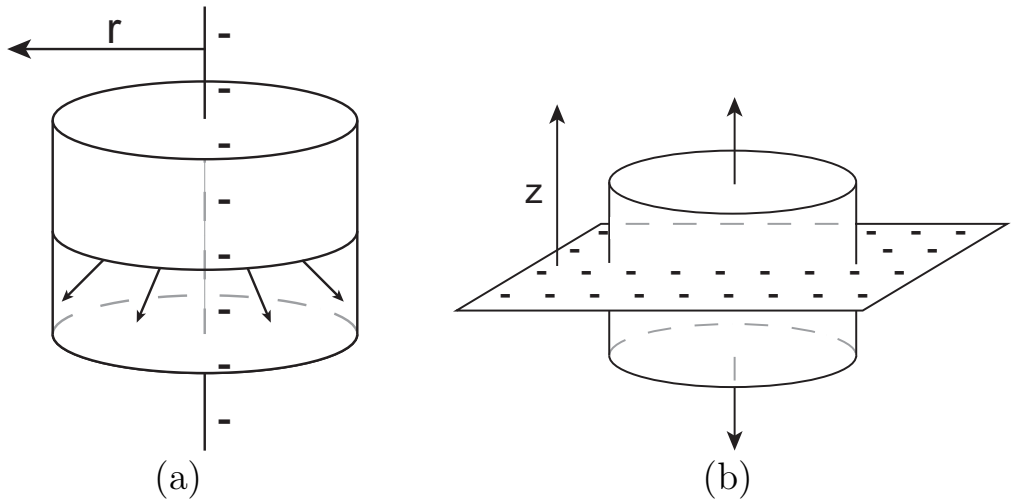


Figure 2.3: (a) Charge distributed along a line, e.g. along DNA. (b) Charge distributed along a plane, e.g. a charged lipid bilayer. In both cases, the fictitious cylinders are used to calculate the electric field (arrows) with the help of the Gauss law.

where ρ is charge density and

$$\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0\epsilon}\int d\mathbf{r}'\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{2.4}$$

is the electrostatic potential. From this, one immediately can verify the Maxwell equations of electrostatics:

$$\nabla\times\mathbf{E} = 0, \qquad \nabla\cdot\mathbf{E} = -\nabla^2\phi = \frac{\rho}{\epsilon_0\epsilon} \tag{2.5}$$

where we have used $\nabla\times\nabla\phi = 0$ and $\nabla^2(1/r) = -4\pi\delta(r)$ with $\delta(r)$ being the Delta-function. The second relation is the Poisson equation. In integral form, it is known as the Gauss law

$$\int_V d\mathbf{r}(\nabla\cdot\mathbf{E}) = \int_{\partial V}\mathbf{E}d\mathbf{A} = \int_V d\mathbf{r}\frac{\rho}{\epsilon_0\epsilon} = \frac{Q_V}{\epsilon_0\epsilon} \tag{2.6}$$

where we have used the divergence theorem to convert the volume into an area integral and the Poisson equation to arrive at Q_V, the overall charge contained in the volume V.

As a first example, we consider the charged line with charge density λ, compare Fig. 2.3a. This would be a model for example for DNA with $\lambda = -2e/0.34$ nm. We now consider a cylinder of radius r and length L around the charged line. For symmetry reasons, the electrostatic field can have a radial component E_r only. Appling the Gauss law Eq. (2.6) gives

$$E_r 2\pi r L = \frac{\lambda L}{\epsilon_0 \epsilon} \;\Rightarrow E_r = \frac{\lambda}{2\pi\epsilon_0\epsilon r} \tag{2.7}$$

Therefore the electrostatic field now scales $\sim 1/r$ rather than $\sim 1/r^2$ as around a point charge. For the potential, we find

$$\phi = -\int^r E_{r'} dr' = -\frac{\lambda}{2\pi\epsilon_0\epsilon} \ln r \tag{2.8}$$

There are two problems with this result, namely the two divergences at small and large distances. The divergence at small distance can be fixed with a molecular cutoff, which gives a constant distribution to the potential and therefore can be neglected. The divergence at large distance is more severe: it implies that it is impossible to fix a finite potential as the boundary condition at infinity. In fact this kind of boundary problem is characteristic for Laplace-type equations in two dimensions. If boundary conditions are an important issue, then a more detailed model is needed.

The next example we consider is the infinitely extended charged plane with charge density σ, compare Fig. 2.3b. This could be a model for example for a bilayer made from charged lipids, with σ of the order of e/nm^2. We now consider a small cylinder crossing the plane, with surface area A at the top and bottom sides. For symmetry reasons, the electrostatic field can have a normal component E_z only. The Gauss law Eq. (2.6) now gives

$$E_z 2A = \frac{\sigma A}{\epsilon_0 \epsilon} \;\Rightarrow E_z = \frac{\sigma}{2\epsilon_0\epsilon},\; \phi = \frac{-\sigma z}{2\epsilon_0\epsilon} \tag{2.9}$$

Therefore the force on a test particle is independent of the distance to the wall and the electrostatic field jumps by $(\sigma/\epsilon_0\epsilon)$ across the surface.

The result Eq. (2.9) has many interesting consequences. For example, consider two equally charged planes facing each other, which might be for example the surfaces of two large and close-by colloids. Eq. (2.9) then predicts that an ion between the two surfaces does not feel any force, irrespective of its position, because the two electric fields are of equal magnitude and point

in opposite directions. Next consider two oppositely charged planes facing each other. Now Eq. (2.9) predicts that an ion outside this arrangement does not feel any force, because again the two electric fields are of equal magnitude and point in opposite directions. This is an extreme example of screening. Between the two planes, an ion would be pulled to the oppositely charged plane.

The latter arrangement is also known as a capacitor in classical electrostatics. An important example for a biological capacitor is the plasma membrane of nerve cells, where ion pumps and channels leads to accumulation of opposite charges on the different sides of the membrane. A capacitor is characterized by the capacitance $C = Q/\Delta\phi$, which describes how much charge Q can be accumulated for a given potential difference $\Delta\phi$. In our case, we have $\Delta\phi = \sigma d/\epsilon_0\epsilon$ and $Q = \sigma A$, resulting in $C/A = \epsilon_0\epsilon/d$. With $d = 2$ nm and $\epsilon = 2$ for the hydrophobic part of a lipid bilayer, we get a value of $\mu F/cm^2$, which is exactly the right order of magnitude as found experimentally for non-myelinated nerve cells. Note that electrostatics at the plasma membrane of neuronal cells is the basis of the way our brain functions (the corresponding field of biology is called *electrophysiology*).

2.2.3 Counterions in solution

We have already observed above that in a stable system, all charges have to be compensated by charges of the opposite sign. Indeed, in soft matter systems charges usually are never far from charges of the opposite sign. The most common situation is a surface immersed in a liquid. There are two mechanisms how such a surface can become charged: dissociable surface groups might release counterions into solution (e.g. protons for acidic groups), and charged particles might get adsorbed from solution (e.g. Ca^{2+} onto the zwitterionic headgroup of lipid bilayers). In both cases, mobile counterions will balance the charges on the surface. While entropy drives them away from the surface, the surface charges keeps them close. In this way, the charged object gets shielded by a diffuse layer of counterions. Together with the surface charges, this gives rise to the so-called *electrical double layer*. The shielding effect becomes stronger when salt is added to the solution, because now there are more counterions being attracted by the surface, while co-ions are repelled. Therefore added salt is an easy way to reduce the electric interaction between charged objects in solution. Famous examples are precipitation of mineral particles when rivers hit the sea (resulting in

river deltas) and compactification of DNA in the cell nucleus by multivalent counterions like spermidine.

Because counterions are mobile, we now for the first time have to deal with thermal averages. Because the particles interact with long-ranged interactions, these seemingly simple systems are very hard problems in theoretical physics. This becomes clearer if we write down the Hamiltonian of this system:

$$\frac{H}{k_B T} = \sum_{i<j} \frac{q^2 e^2}{4\pi\epsilon_0 \epsilon k_B T r_{ij}} + \sum_i \frac{q e^2 \sigma z_i}{2\epsilon_0 \epsilon k_B T} = \sum_{i<j} \frac{q^2 l_B}{r_{ij}} + \sum_i \frac{z_i}{\mu} \tag{2.10}$$

where q now denotes valency (that is the charge is qe) and σ is the area number density of elementary charges (that is the area charge density is $e\sigma$). Here we have used the result Eq. (2.9) for a charged plane. Eq. (2.10) introduces two new quantities. The *Bjerrum length* $l_B = e^2/(4\pi\epsilon_0\epsilon k_B T)$ is the distance at which two unit charges interact with thermal energy. In vacuum, its value is 56 nm, while in water, it goes down to 0.7 nm. The *Gouy-Chapman length* $\mu = 1/(2\pi q\sigma l_B)$ is the distance from the wall at which the potential energy equals thermal energy. It can also be identified with the thickness of the diffusive counterion layer. For a number density at the charged wall of $\sigma = 1/nm^2$, its value is 1 nm (for $q = 1$). We now rescale all distances with μ to get a dimensionless Hamiltonian

$$\frac{H}{k_B T} = \sum_{i<j} \frac{\Xi}{\bar{r}_{ij}} + \sum_i \bar{z}_i \tag{2.11}$$

where we have introduced the coupling strength

$$\Xi = \frac{l_B q^2}{\mu} = 2\pi q^3 l_B^2 \sigma = \frac{q^3 e^4 \sigma}{8\pi(\epsilon\epsilon_0 k_B T)^2} \tag{2.12}$$

Note that the high powers in q and T results from the different scaling of point-point and point-plane interactions, i.e. it is a geometrical effect.

Eq. (2.12) shows that there are essentially two regimes here. For low charge or high temperature one has $\Xi < 1$, correlation effects are small and one can use a mean field theory, the so-called *Poisson-Boltzmann theory*. For added salt, this theory can be further simplified by linearization, leading to *Debye-Hückel theory*. For high charge or low temperature one has $\Xi > 1$, the so-called *strong coupling limit*. Then correlation effects become important and can be treated with a field theoretical virial expansion. Eq. (2.12)

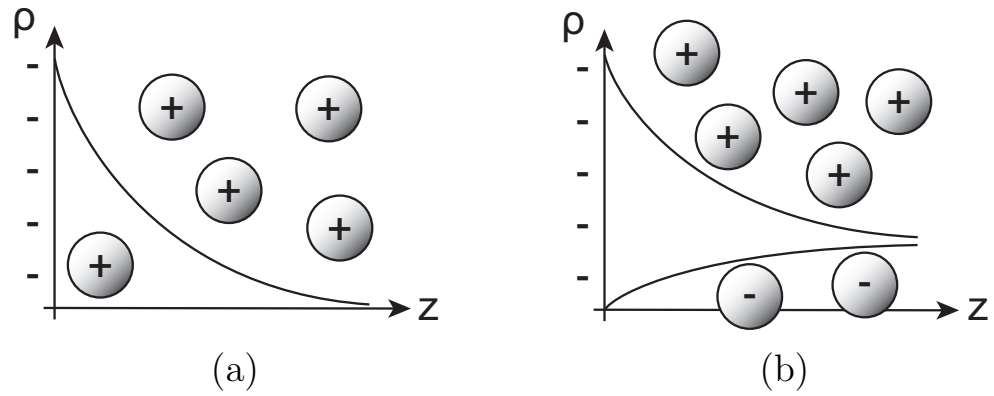


Figure 2.4: (a) Counterions in solution form a diffusive layer around a charged surface, resulting in the electrical double layer. Charge condensation is counteracted by entropy. (b) For added salt, coions are repeled from the surface and the electrical double layer is enhanced. Therefore salt increases screening effects.

also shows that valency q is a crucial parameter. In fact it is well known in biology that condensation problems (like DNA compactification) often involves multivalent ions. It is also important to note that the strong coupling limit indeed is experimentally accessible: with trivalent ions ($q = 3$) and $\sigma = 1/nm^2$, we already have $\Xi = 100$.

2.2.4 One charged surface

We first consider one charged surface, compare Fig. 2.4a. The basic idea here is self-consistency between the potential and the charge density. The potential has to satisfy the Poisson equation $\nabla^2\phi = -\rho/\epsilon_0\epsilon$. For the volume charge density ρ, we neglect correlation effects and assume that it follows as for a single particle in an external field. In thermal equilibrium, the number density n follows from a Boltzmann distribution, $n = n_0 \exp(-e\phi/k_BT)$ (for $q = 1$). Moreover $\rho = en$. Combining these equations gives the *Poisson-Boltzmann equation*

$$\nabla^2\phi = -\frac{e}{\epsilon_0\epsilon} n_0 e^{-\frac{e\phi}{k_BT}} \tag{2.13}$$

which is a second order non-linear differential equation for ϕ. Two boundary conditions are required for its solution. The first boundary condition is that no force acts at infinity, i.e. $\phi'(\infty) = 0$. A second boundary condition follows

from charge neutrality:

$$\sigma = -\int_0^\infty \rho dz = \epsilon_0\epsilon \int_0^\infty \phi'' dz = -\epsilon_0\epsilon\phi'(0) \Rightarrow \phi'(0) = \frac{\sigma}{\epsilon_0\epsilon} \tag{2.14}$$

With these two boundary conditions, Eq. (2.13) can be solved analytically:

$$\phi = \frac{2k_BT}{e}\ln(z+\mu)\ , \qquad n = \frac{1}{2\pi l_B}\frac{1}{(z+\mu)^2} \tag{2.15}$$

In contrast to the result $\phi \sim z$ for the case without counterions, now the electrostatic potential scales only $\sim \ln z$. The counterion profile shows a long-ranged (power-law) decay.

Next we consider the case of added salt, compare Fig. 2.4b. For simplicity, we consider a simple 1:1 electrolyte like sodium chloride. Moreover we note that far away from the wall, charge is carried only by the salt with concentration n_0. We therefore write

$$\nabla^2\phi = -\frac{e}{\epsilon_0\epsilon}n_0(e^{-\frac{e\phi}{k_BT}} - e^{\frac{e\phi}{k_BT}}) = \frac{2e}{\epsilon_0\epsilon}n_0 \sinh\frac{e\phi}{k_BT} \approx \kappa^2\phi \tag{2.16}$$

The linearization is valid for small potential or high temperature and the linearized Poisson-Boltzmann equation is known as the *Debye-Hückel equation*. In Eq. (2.16) we have introduced a new length scale, the Debye screening length $1/\kappa = (\epsilon_0\epsilon k_BT/2e^2q^2n_0)^{1/2} = 1/(2n_0l_B)^{1/2}$. Its values are 1 μm, 10 nm, 1 nm and 0.3 nm for 10^{-7} M (pure water), 10 mM, 100 mM (physiological salt concentration) and 1 M, respectively. For the planar geometry, Eq. (2.16) is solved by $\phi = (\sigma/\epsilon_0\epsilon\kappa)e^{-\kappa z}$, i.e. the potential decays exponentially fast on the scale of the Debye screening length. Thus the decay is long-ranged in Poisson-Boltzmann theory (without salt) and short-ranged in Debye-Hückel theory (with salt).

For a sphere with radius R, which could be a protein, ion or micelle, we have to solve

$$\nabla^2\phi = \frac{1}{r}\frac{d^2(r\phi)}{dr^2} = \kappa^2\phi \ \Rightarrow \phi = \frac{R\phi_R}{r}e^{-\kappa(r-R)} \tag{2.17}$$

The surface potential ϕ_R follows from the Gauss law Eq. (2.6):

$$E_R = \frac{Q_V}{4\pi\epsilon_0\epsilon R^2} = -\phi'|_R = \frac{\phi_R(1+\kappa R)}{R} \ \Rightarrow \phi_R = \frac{Q_V}{4\pi\epsilon_0\epsilon(1+\kappa R)R} \tag{2.18}$$

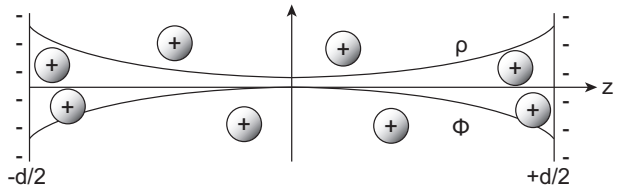


Figure 2.5: If two charged surfaces oppose each other, both the counterion distribution n and the electrostatic potential ϕ are symmetric around the origin. Counterion entropy leads to a disjoining pressure between the surfaces.

There are two important special cases of Eq. (2.18). For no salt, $\kappa = 0$ and $\phi = Q_V/(4\pi\epsilon_0\epsilon r)$, thus we recover the Coulomb law. For a point charge, $R = 0$ and $\phi = Q_V e^{-\kappa r}/(4\pi\epsilon_0\epsilon r)$ (*Yukawa-potential*). Therefore the interaction between point-like charges in the case of salt becomes very short-ranged.

Since we deal with a linear theory, total energy follows by simple summation (in other words, the Yukawa-potential is the Green function for Debye-Hückel theory). For charged and dilute colloids in solution, the time scales of particle and counterion movement and the length scales of particle size and separation are usually sufficiently well separated such that a statistical treatment can be based on effective interaction laws like the Yukawa-potential. It also can be used to derive the potential around spatially distributed charges which are screened by salt. By integrating the Yukawa-potential for a charged line with charge line density λ, we find $\phi = 2k_BTl_B\lambda K_0(\kappa r)$, where K_0 is the Bessel function. For $\kappa \to 0$ (salt-free limit), we retrieve the result $\phi = -\lambda \ln r/2\pi\epsilon_0\epsilon$, compare Eq. (2.8). By integrating the Yukawa-potential for a charged plane with charge area density σ, we get $\phi = 2\pi k_BTl_B\sigma e^{-\kappa z}/\kappa$, which for $\kappa \to 0$ results in $\phi = -\sigma z/2\epsilon_0\epsilon$, compare Eq. (2.9).

In regard to the strong coupling limit, we only report the result to first order in the virial expansion, $n = 2\pi l_B\sigma^2 e^{-z/\mu}$. Although an exponential decay occurs like in Debye-Hückel theory, this result has nothing to do with linearization of the Poisson-Boltzmann theory. In particular, here the decay occurs over the Gouy-Chapman length μ rather than over the Debye screening length $1/\kappa$.

2.2.5 Two charged surfaces

We now consider the distribution of counterions between two similarly charged surfaces with distance d, which we take to be centered around the origin,

compare Fig. 2.5. This situation is a model e.g. for two colloids at small separation. Experimentally it can be investigated with a setup called the *surface force apparatus*. For the case without salt, we again have to solve the Poisson-Boltzmann equation Eq. (2.13), only that now the boundary conditions are slightly different. For symmetry reasons, the first boundary condition is simply $\phi'(0) = 0$. The second boundary condition again follows from charge neutrality:

$$\sigma = -\int_0^{d/2} \rho dz = \epsilon_0 \epsilon \int_0^{d/2} \phi'' dz = \epsilon_0 \epsilon \phi'(d/2) \Rightarrow \phi'(d/2) = \frac{\sigma}{\epsilon_0 \epsilon} \tag{2.19}$$

Then the exact solution is

$$\phi = \frac{k_B T}{e} \ln(\cos^2(Kz)) \ , \qquad n = \frac{n_0}{\cos^2(Kz)} \tag{2.20}$$

where n_0 now is the counterion density at the midplane and K is a constant given by $K^2 = (e^2 n_0)/(2\epsilon_0 \epsilon k_B T)$. Thus once K is known, n_0 follows as $n_0 = (2K^2 \epsilon_0 \epsilon k_B T)/e^2$. K has to be determined as a function of σ and d:

$$\phi'(d/2) = \frac{2k_B T K}{e} \tan \frac{Kd}{2} = \frac{\sigma}{\epsilon_0 \epsilon} \tag{2.21}$$

Eq. (2.20) shows that the counterion profile between two similarly charged surfaces is sort of parabolic, with counterions accumulating at the sides. Superficially this reminds of the situation with one surface, when counterions were attracted to the oppositely charged surface. However, the situation now is fundamentally different, because as we have seen above for a single charge in front of a charged surface, the electrostatic force on a single ion is independent of distance and the two forces from the two surfaces cancel each other. In fact the real reason why counterions accumulate at the sides is mutual repulsion between different counterions. This situation is similar to the accumulation of mobile charges on the surface of any conductor with mobile charges. Charge accumulation is further favored by the fact that negative and positive charges can condense into stable structures, as we have seen for the NaCl-crystal. So why do the counterions not simply condense onto the surfaces ? The answer is that we deal with a high temperature situation and configurational entropy drives them back into solution, therefore entropy is an essential element to understand Eq. (2.20). In summary, it is essentially the counterion pressure (*disjoining pressure*) which pushes the surfaces away from each other.

The exact value for the the disjoining pressure follows in a very general way from the *contact value theorem*, which is stated here without derivation:

$$p = k_B T n_0 \tag{2.22}$$

Thus the pressure on the surface is simply the osmotic counterion pressure at the midplane. Using the Poisson-Boltzmann result for n_0, Eq. (2.22) gives

$$p = 2\epsilon_0 \epsilon \left(\frac{k_B T}{e}\right)^2 K^2 \tag{2.23}$$

This equation has two interesting features: first it scales $\sim T^2$, indicating that we deal with an entropic effect, which vanishes with temperature T, and second we see that $p > 0$, that is the pressure is always repulsive. For large separation, $d \to \infty$ and therefore Eq. (2.21) demands that in order to give a finite value for the charge density σ, we must have $Kd/2 \approx \pi/2$. Therefore

$$p \approx 2\epsilon_0 \epsilon \left(\frac{k_B T}{e}\right)^2 \left(\frac{\pi}{d}\right)^2 \tag{2.24}$$

The pressure in Poisson-Boltzmann theory is a long-ranged effect.

It is instructive to put numbers on these results. For $\sigma = 0.2$ C/m^2 (roughly one elementary charge per nm^2) and $d = 2$ nm, we find $K = 1.3 \times 10^9$ 1/m. Therefore the counterion concentration at the midplane is $n_0 = 0.7$ M. At the surface, it is $n(d/2) = 12$ M. Therefore it is a factor 18.5 higher at the sides than in the middle, and this over a distance of only 1 nm. For the electrostatic potential, we have $\phi(0) = 0$ by construction and $\phi(d/2) = 74$ mV at the surface. The pressure between the surfaces follows as $p = 17$ atm, which is much larger than atmospheric pressure. An important biological application of this large osmotic counterion pressure is the way the cartilage surfaces in our knee joints are prevented from direct contact. In this case, the counterions balance the charges on certain polymers attached to the surfaces, but the fundamental mechanism is the same. Thus counterion pressure is the reason why we can go jogging.

We only briefly discuss the two other interesting regimes of Debye-Hückel theory and the strong coupling limit. In the first case of added salt, the exact solution reads $\phi = \phi_0 \cosh(\kappa z)$. Thus the concentration profile and the pressure decay exponentially with distance z, that is they are short-ranged. In the strong coupling limit, entropic effects become less relevant and

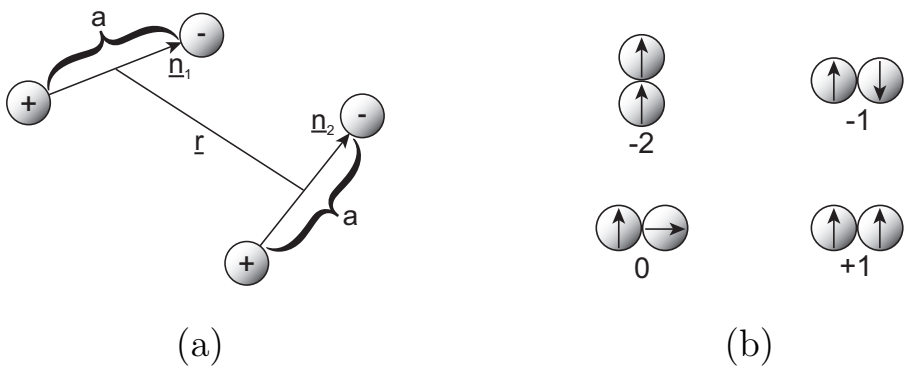


Figure 2.6: (a) The interaction between two dipoles is strongly orientation-dependent. The angles involved follow from the two orientation vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ and the normalized separation vector $\mathbf{r}$. (b) The most favorable configuration is parallel alignment, with all other possible configurations being less favorable. Therefore dipolar fluids are characterized by chaining.

charges tend to condense onto the surfaces. A detailed analysis shows that for large coupling strength Ξ, attraction between similarly charged surfaces becomes possible, that is pressure becomes negative. Although this result sounds surprising, it is well known experimentally that highly charged planar surfaces attract each other in the presence of multivalent counterions. For example, this effect restricts the swelling of e.g. clay or lipid lamellar phases. Strong coupling theory also explains why multivalent ions are so prominent in compactification of highly charged biological material like DNA.

2.3 Dipolar and van der Waals interactions

Up to now we have considered the long-ranged Coulomb-interaction. We have found that geometry modulates the interaction laws and that screening by mobile counterions can considerably shorten their range. Another mechanism by which short-ranged interactions arise in soft matter systems is the electrostatic interaction of electric dipoles. These dipoles can be either permanent (like in CO, H_2O, NH_3) or they can result transiently from charge fluctuations.

The interaction between two dipoles follows from a multipolar expansion

of the Coulomb potential as

$$U = \frac{(ea)^2}{\epsilon_0 \epsilon r^3} \left[(\mathbf{n}_1 \cdot \mathbf{n}_2) - 3(\mathbf{n}_1 \cdot \mathbf{r})(\mathbf{n}_2 \cdot \mathbf{r}) \right] \tag{2.25}$$

where a is the distance between the two charges in a dipole, $\mathbf{r}$ is the unit vector between the two dipoles and the $\mathbf{n}_i$ are the unit vectors describing the two dipole orientations. Fig. 2.6a depicts these definitions. In Eq. (2.25), the term in brackets describes the angular part of the interaction. Its evaluation in Fig. 2.6b shows that parallel alignment in a chain is the most favorable configuration. In regard to distance, the dipolar interactions decays $\sim 1/r^3$. As the charges can reorient, it is generically attractive. Dipolar interactions can be conveniently studied in colloidal systems, namely with *dipolar fluids*, which are characterized by disordered networks of chains at low density and by spontaneous polarization at high density.

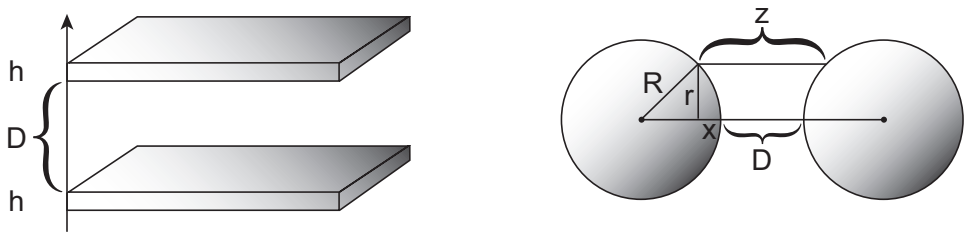
When interaction energies are low, e.g. in water with its high dielectric constant ϵ, then the dipoles are subject to strong configurational fluctuations and one has to perform a thermal average over all possible orientations. An effective interaction W can then be defined by

$$e^{-W/k_BT} = \langle e^{-U/k_BT} \rangle = \frac{\int e^{-U/k_BT} d\Omega}{\int d\Omega} \tag{2.26}$$

Since both W and U are small compared with thermal energy k_BT, one next can expand the two exponentials. For Eq. (2.25), the first order term vanishes and the second order term gives the so-called *Keesom interaction*

$$W = -\frac{\langle U^2 \rangle}{2k_BT} = -\frac{(ea)^4}{3(4\pi\epsilon_0\epsilon)^2 k_BT r^6} \tag{2.27}$$

which generically is attractive. Similarly short-ranged interactions also result from polarization effects. A permanent or transient dipole gives an electric field $E \sim 1/r^3$, which in a polarizable medium leads to a new dipole $p \sim E$, thus resulting in $W = -pE \sim -1/r^6$. For permanent and transient dipoles, these are the so-called *Debye and London interactions*, respectively, with the latter being an example of a dispersion force resulting from quantum mechanical fluctuations. Since the Keesom, Debye and London interactions all scale as $\sim -1/r^6$, the are collectively called *van der Waals interaction*. It is the main source of cohesion in soft condensed matter systems.



(a) (b)

Figure 2.7: (a) Two films of thickness h interacting over a distance D. (b) Two spheres of radius R interacting over a distance D. For close approach, $D \ll R$, the interaction can be calculated from the Derjaguin approximation.

Like the Coulomb interaction, the van der Waals interaction is strongly modulated by geometry. In many cases, one can assume pairwise additivity (*Hamacker approach*):

$$W = -\frac{A}{\pi^2} \int_{V_1} d\mathbf{r}_1 \int_{V_2} d\mathbf{r}_2 \frac{1}{{r_{12}}^6} \tag{2.28}$$

The Hamacker constant A has a typical value of 25 $k_B T$. For two bodies made from material 1 and 2 in a medium m it scales like $\sim (\epsilon_1 - \epsilon_m)(\epsilon_2 - \epsilon_m)$. Therefore $A > 0$ for $1 = 2$, that is like bodies attract. However, for $\epsilon_1 < \epsilon_m < \epsilon_2$, van der Waals interaction can also leads to repulsion.

A large variety of different geometries have been treated in the Hamaker approach, including the interaction between thin and thick films, halfspaces, hollow and full spheres, cylinders and strings. As an instructive example, here we discuss in detail the interaction between two thick films, each of thickness h, and separated by a distance D, compare Fig. 2.7a. This could be a model for the van der Waals-interaction between two lipid bilayers in a lamellar stack. First we integrate over one film:

$$W_1(D) = -\frac{A}{\pi^2} 2\pi \int_0^\infty r dr \int_D^{D+h} dz \frac{1}{(r^2 + z^2)^3} \tag{2.29}$$

$$= -\frac{A}{2\pi} \int_D^{D+h} dz \frac{1}{z^4} = -\frac{A}{6\pi} \left[\frac{1}{D^3} - \frac{1}{(D+h)^3} \right] \tag{2.30}$$

Next we integrate over the second film. Due to translational invariance in the x-y-plane, this results in an energy per area:

$$w(D) = \int_D^{D+h} dD'\ W_1(D') = -\frac{A}{6\pi}\frac{h^2(3D^2+6Dh+2h^2)}{D^2(D+h)^2(D+2h)^2} \tag{2.31}$$

$$= -\frac{A}{12\pi}\left[\frac{1}{D^2} - \frac{2}{(D+h)^2} + \frac{1}{(D+2h)^2}\right] \tag{2.32}$$

The two equivalent expressions are useful to consider two limits of interest. In the case of thin films, $h \to 0$ and one has

$$w(D) = -\frac{Ah^2}{2\pi D^4} \tag{2.33}$$

Note that the scaling part of this result can easily be guessed: in the limit of small thickness, w has to scale $\sim h^2$, and D is the only other length scale present to obtain an energy per area. In the case of halfspaces, $D \to 0$ and one has

$$w(D) = -\frac{A}{12\pi D^2} \tag{2.34}$$

Again it is easy to guess the scaling form, but obviously the full calculation is required to get the numerical prefactors, which here amount to a more than one order of magnitude reduction. Because there are only two length scales in this problem, the crossover between the two cases occurs at $D \approx h$. As in the electrostatic case, again we note that different geometries lead to different scaling laws. In particular, the two halfspaces show a van der Waals-interaction which is much more long-ranged ($\sim 1/D^2$) than the ones between two thin films ($\sim 1/D^4$) or single molecules ($\sim 1/r^6$).

The integrals of the Hamacker approach can also be solved exactly for the case of full and hollow spheres, which are important models for colloids. For example, the result for two full spheres with radii R_1 and R_2, respectively, and distance r reads

$$W_{R_1R_2} = -\frac{A}{6}\left\{\frac{2R_1R_2}{r^2-(R_1+R_2)^2} + \frac{2R_1R_2}{r^2-(R_1-R_2)^2} + \ln\left(\frac{r^2-(R_1+R_2)^2}{r^2-(R_1-R_2)^2}\right)\right\} \tag{2.35}$$

The result can also be used to obtain the interaction between two hollow spheres with finite thickness:

$$W_{S_1S_2} = W_{B_1B_2} - W_{C_1C_2} - W_{C_1S_2} - W_{S_1C_2} = W_{B_1B_2} + W_{C_1C_2} - W_{C_1B_2} - W_{B_1C_2} \tag{2.36}$$

where we have decomposed each sphere into ball (B) and core (C) and repeatingly used $W_{CB} = W_{CS} + W_{CC}$. The remaining expressions can now be evaluated using Eq. (2.35).

In many cases one is interested in the interaction between two objects at small separation. In the case of two spheres, using $R_1 = R_2 = R$ and $D = r - 2R \ll R$ in Eq. (2.35) leads to

$$W = -\frac{AR}{12D} \tag{2.37}$$

Thus the decay is slower than between two halfspace, because less material is moved away upon separation. Another and more general way to obtain this result is the use of the *Derjaguin approximation*, which is a perturbation analysis around the planar case. The basic idea is that for a given point on one surface, one determines the point of closest approach z on the other surface. One then integrates $w(z)$ for the interaction between planar surfaces over the first surface. Because for the two spheres the problem has radial symmetry, we have

$$W = 2\pi \int_0^\infty w(z) r dr \tag{2.38}$$

The geometrical situation is depicted in Fig. 2.7b. For large R, the relation between z and r follows as $z \approx D + r^2/R$. Therefore

$$W = \pi R \int_D^\infty w(z) dz \tag{2.39}$$

In principle, this expression is valid for any type of interaction $w(z)$. Using Eq. (2.34) for the van der Waals-interaction between two halfspaces, we find for the van der Waals-interaction between two spheres at close approach:

$$W = -\frac{AR}{12} \int_D^\infty \frac{1}{z^2} dz = -\frac{AR}{12D} \tag{2.40}$$

in agreement with Eq. (2.37). Another useful result follows if we use Eq. (2.39) to obtain the force between the two spheres:

$$F = -\frac{\partial W}{\partial D} = \pi R w(D) \Rightarrow w(D) = \frac{F}{\pi R} \tag{2.41}$$

Therefore measuring the force for any kind of interaction between two spheres immediately gives the corresponding interaction law between two flat surfaces. A related result for two crossed cylinders is the starting point for work with the surface force apparatus.

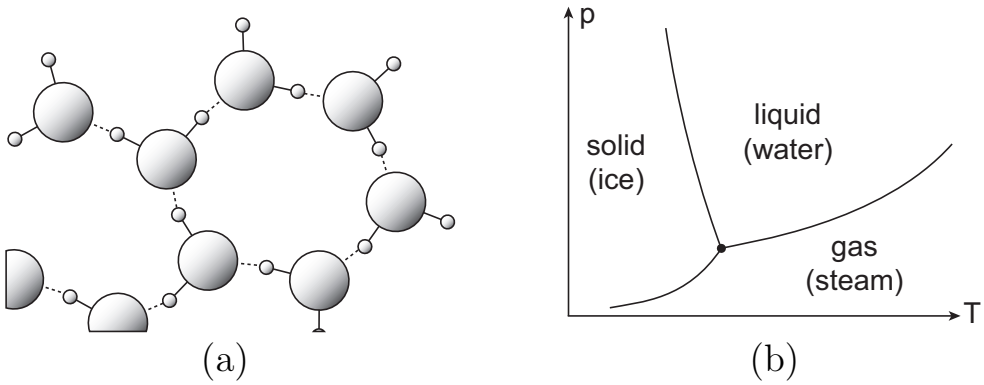


Figure 2.8: (a) Water can be regarded as a fluctuating network of orientation-dependent hydrogen bonds (broken lines). (b) The phase diagram of water as a function of temperature T and pressure p has an unusual negative slope of the solid-liquid boundary, reflecting that for water, the liquid is denser than the solid.

2.4 Hydrophilic and hydrophobic interactions

Much of the complexity of soft matter and biological systems arises from the peculiar properties of water. Unlike most other solvents, water should not be regarded as an ensemble of identical spheres interacting through a simple pair potential, but rather as a fluctuating network of hydrogen bonds, compare Fig. 2.8a. The dominance of the hydrogen bonds in water arises because the two O-H-bonds in a water molecule are proton donors, while the two electron pairs are acceptors. Therefore four hydrogen bonds arise around one water molecule in bulk water, with the energetically most favorable arrangement being tetrahedral. In general, the structure of water is determined by the tendency to maximize the number of hydrogen bonds. This tendency weakens as temperature increases because the systems becomes more randomized due to entropy. Like for other hydrogen-bonded liquids (HF, H_2O_2, HCN), changes in electron density quickly propagate throughout the solvent, leading to strong cooperativity.

The picture of water as a network of hydrogen bonds helps to understand some of its peculiar properties:

- The high dielectric constant of $\epsilon = 80$ arises because the hydrogen bonds are easily polarized.
- The high cohesive energy stored in the network of hydrogen bonds

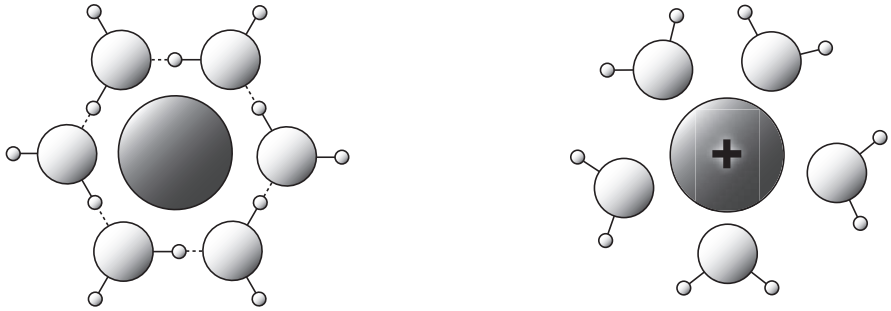


Figure 2.9: Water is ordered in the presence of solute. (a) Water cannot establish hydrogen bonds to a hydrophobic (oil-like) solute, so effectively it avoids it. (b) Ions can enhance or weaken the structure of the surrounding water, depending on molecular details and temperature.

results in high values for the material properties describing cohesion, including melting temperature, boiling temperature, enthalpy of vaporization, surface tension and heat capacity. In general, there is a competition between hydrogen bonds favoring open structures and van der Waals bonds favoring dense structures, which is strongly modulated by temperature.

- The solid is less dense than the liquid (ice floats on water) because the tetrahedral ordering results in an open, almost foam-like structure, which is destroyed by temperature. This effect persists even in the liquid, therefore heating increases density up to 4 degrees Celsius (above 4 degrees Celsius, entropy leads to the usual effect of thermal expansion, therefore density has a maximum at 4 degrees Celsius). Because pressure like heating effectively destroy the open structure of ice, it can be used to melt the crystal, in contrast to most other solvents (this is one of the reasons why one can skate on ice but not on glass). Therefore the solid-liquid boundary has an unusual negative slope in the p-T-diagram, compare Fig. 2.8b.

The hydrogen bonds of water are also responsible for its special properties as solvent. The most prominent effect here is the *hydrophobic effect*, which means that oil and water do not mix. The main mechanism here is that oil molecules immersed in water do not participate in the network

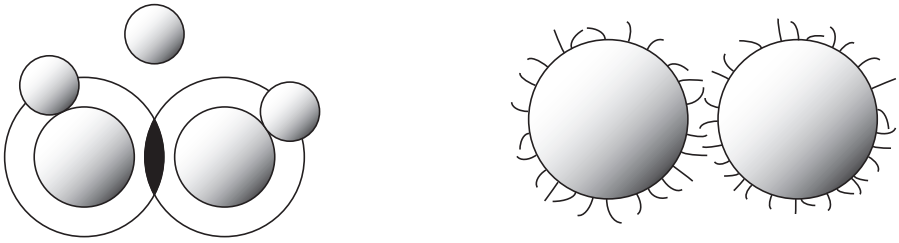
of hydrogen bonds. Locally, this leads to a higher order around the solute, which is entropically unfavorable, as shown schematically in Fig. 2.9a. Indeed calorimetry shows that the aversion of oil for water is entropic at 25 degrees Celsius (it becomes enthalpic at higher temperatures). Closely related is the issue of *hydrophobic interactions*, which means that non-polar solutes aggregate in water as if there was some attractive interaction, because in this way the solutes reduce the surface area of contact with the network of hydrogen bonds. For polar or ionic solutes, the situation is more complicated, that is depending on the details, they can be favorable or unfavorable for the structure of water. This is shown schematically in Fig. 2.9b. The term *hydrophilic interactions* refers to the fact that polar solutes prefer polar solvents since a large dielectric constant lowers the self-energy.

2.5 Colloidal dispersions

Colloidal dispersions are mesoscopic particles in the size range from 10 nm to 10 μm which are dispersed in solvent. For example, milk is a colloidal dispersion of fat droplets in water and it is opaque because the droplet have a size similar to optical wavelengths. In general, colloidal particles tend to flocculate due to the ubiquitous van der Waals interaction. Therefore colloidal dispersions have to be stabilzed. The two common mechanisms to do so are electrostatic or steric repulsion between charged and polymer-decorated surfaces, respectively.

Steric effects can lead to both attraction and repulsion between colloids, compare Fig. 2.10. Due to the depletion interaction, large particles (e.g. latex beads) in a sea of small particles (e.g. coiled polymers) attract each other. The reason is that each large particle is surrounded by a zone which is excluded to the centers of mass of the small particles due to their finite size, compare Fig. 2.10a. If two large particles aggregate, the excluded volume decreases due to overlap of these zones, which is favorable for entropic reasons. This implies that the interaction range is basically set by the size of the small particles. Repulsion results between two colloids which are surface-grafted with polymers. As the two surfaces approach, the polymers start to overlap, compare Fig. 2.10b. This is unfavorable for entropic reasons, leading to a repulsive force between the two surfaces.

The model systems for colloidal physics are polystyrene (PS) or polymethylmethacrylate (PMMA) spheres dispersed in water. Since the polymer



(a) (b)

Figure 2.10: Colloidal interactions based on steric effects. (a) Depletion interaction: small particles define exclusion zones around large particles. Overlap between these exclusion zones (black) is favorable because it effectively increases the free volume for the small particles. (b) Steric stabilization: polymers grafted to the surfaces of the colloids prevent coagulation because overlap is sterically unfavorable.

brushes are much smaller than particle sizes and since the van der Waals interaction can be suppressed by index matching between particles and solvent (i.e. the dielectric constants are made equal by addition of salt), sterically stabilized PMMA-spheres are a model system for hard spheres. Charge stabilization requires some surface groups which dissociate counterions into the solution, which is the case for e.g. PS-spheres in water. Typically there are 1000 elementary charges on such a macroion. Today the interactions in colloidal systems can be tuned to such a degree that they can be used as model systems for the study of atomic and molecular systems.

The combined action of van der Waals attraction and double-layer repulsion is described by the so-called *DLVO-theory*, named after Derjaguin, Landau, Verwey and Overbek. Since the van der Waals-interaction diverges with a power law at small separation, while the electrostatic interaction stays finite, a primary minimum results for small separation. This leads to coagulation and flocculation, e.g. at high salt when electostatic repulsion is screened (an example mentioned before is precipitation of mineral particles in river deltas). However, at low salt the electrostatic repulsion can dominate at intermediate distances, leading to an energy barrier which provides kinetic stability because particles become trapped in a secondary minimum. At large separation, van der Waals dominates again. DLVO-theory has been nicely

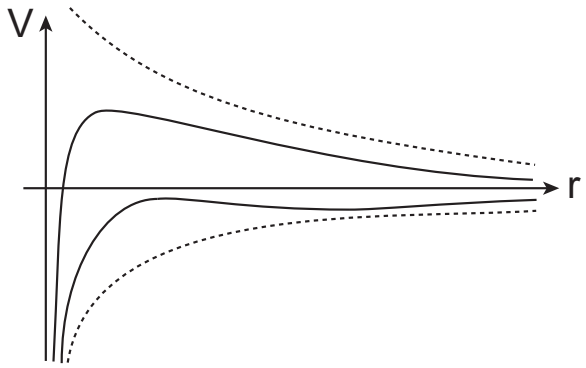


Figure 2.11: DLVO-theory describes the interplay between the attractive van der Waals-interactions (lower curve) and the repulsive screened electrostatic interactions between similarly charges colloids (upper curve). At very small separation, the van der Waals-interaction always wins, leading to the primary minimum. For certain parameter values (lower solid line), a barrier can arise at intermediate distances, resulting in a secondary minimum.

confirmed with experiments with the surface force apparatus.

Chapter 3

Simple and complex fluids

In this chapter, we will discuss the phase behaviour of simple and complex fluids. We start with a review of the statistical mechanics of the ideal gas, which is a good model for diluted gases. In order to deal with real gases, we use the ideal gas as a reference state for a virial expansion. Already the first order of this expansion is sufficient to predict the fluid-fluid phase transition resulting for attractive interactions and the fluid-solid phase transition resulting from excluded volume effects. Both transitions together explain the phase behaviour of Lennard-Jones and square well systems. Finally we will discuss the phase behaviour of more complex fluids such as multicomponent fluids and liquid crystals.

3.1 Ideal gas

The ideal gas is a collection of N non-interacting point particles. It is one of the most important examples for an exactly solvable statistical mechanics model and its understanding provides the background for all work on simple and complex fluids. We start with the partition sum

$$Z = \frac{1}{N!} \int \frac{d^N \mathbf{p} d^N \mathbf{r}}{h^{3N}} e^{-\beta H(\mathbf{p},\mathbf{r})} = \frac{z^N}{N!} \tag{3.1}$$

where $H = \sum_i \mathbf{p}_i^2/2m$ is the ideal gas Hamiltonian (only kinetic energy), $\mathbf{p}_i$ and $\mathbf{r}_i$ are momenta and positions, respectively, of the different particles ($1 \leq i \leq N$). $\beta = 1/k_B T$ and h is Planck's constant. It enters here because the different possible states are assumed to be squeezed together in phase

space as closely as permitted by Heisenberg's uncertainty principle, $\Delta p \Delta q \geq h$. The factor $N!$ accounts for the indistinguishability of the particles. z is the partition sum for one particle:

$$z = \int \frac{d\mathbf{p}d\mathbf{r}}{h^3} e^{-\beta \frac{\mathbf{p}^2}{2m}} = \frac{V}{h^3} \left(2\pi k_B T m\right)^{3/2} = \frac{V}{\lambda^3} \tag{3.2}$$

where we have evaluated the Gauss integral $\int dx e^{-ax^2} = (\pi/a)^{1/2}$ and

$$\lambda = \sqrt{\frac{h^2}{2\pi m k_B T}} \tag{3.3}$$

is the *thermal (de Broglie) wavelength*. Then the free energy follows with the help of Stirling's formula $\ln N! \approx N \ln N - N$ for large N as

$$F = -k_B T \ln Z = -k_B T \ln \left(\frac{z^N}{N!}\right) = -k_B T N \left(\ln \left(\frac{V}{\lambda^3 N}\right) + 1\right) \tag{3.4}$$

From the result for the free energy $F = F(N, V, T)$, we can now calculate the pressure p as

$$p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} = k_B T \frac{N}{V} \Rightarrow pV = N k_B T \tag{3.5}$$

The result is known as the *thermal equation of state* or simply as the *ideal gas law*. It can be understood in simple terms if one considers the momentum transfer onto the wall of a box of dimension L^3 per time and per area:

$$p = N(2mv_x)(\frac{v_x}{2L})\frac{1}{L^2} = mv_x^2 \frac{N}{V} = k_B T \frac{N}{V} \tag{3.6}$$

where he have used the fact that there is only kinetic energy which is directly related to thermal energy through the thermal equation of state.

Next we calculate entropy S:

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} = -\frac{F}{T} + \frac{3}{2} k_B N = k_B N \left(\ln \left(\frac{V}{\lambda^3 N}\right) + \frac{5}{2}\right) \tag{3.7}$$

Then the internal energy U follows as

$$U = F + TS = F + T(-\frac{F}{T} + \frac{3}{2} k_B N) = \frac{3}{2} N k_B T \tag{3.8}$$

This result is called the *caloric equation of state*. It states that the internal energy of an ideal gas only depends on temperature. It is also an example for the equipartition theorem: every quadratic term in the Hamiltonian contributes a term $k_BT/2$ to the internal energy, and here we have 3 translational degrees of freedom.

Historically the thermal and caloric equations of state for the ideal gas have first been found empirically for weakly interacting particles like He, Ar, Ne, O_2, NO or for real gases at low densities. In terms of density $\rho = N/V$, we can write our main results in a very compact way:

$$f = \frac{F}{V} = k_BT\rho\left(\ln(\rho\lambda^3) - 1\right) \tag{3.9}$$

$$p = \rho k_BT \tag{3.10}$$

$$u = \frac{U}{V} = \frac{3}{2}\rho k_BT \tag{3.11}$$

$$\tag{3.12}$$

3.2 Virial expansion for real gas

The ideal gas is a non-interacting system and can be regarded as the small density limit for real gases, in which the components interact with some potential U. We now build up a perturbative scheme to treat density corrections to the ideal gas. If momenta enter the Hamiltonian only through the kinetic energy as above (that is if potential energy does not depend on momenta), then the classical canonical partition function separates into an ideal gas part and a configurational part:

$$Z = Z_{ideal}Z_{inter} \tag{3.13}$$

where

$$Z_{ideal} = \frac{V^N}{N!\lambda^{3N}} \tag{3.14}$$

as above and

$$Z_{inter} = \frac{1}{V^N}\int d^N\mathbf{r}e^{-\beta U(\mathbf{r})} \tag{3.15}$$

This term does not factor into single particle functions because the potential U couples all coordinates. Yet all thermodynamic quantities separate into

an ideal gas part and a correction due to the interactions. In particular, we have

$$F = -k_B T \ln Z = F_{ideal} + F_{inter} \tag{3.16}$$

$$p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} = p_{ideal} + p_{inter} \tag{3.17}$$

The formulae for the ideal expressions have been given above and are valid for dilute fluids. For the pressure, one expects that the correction terms should scale at least in second order in ρ, because two particles have to meet in order to give a contribution to this term. This suggests to make the following ansatz of a Taylor expansion in ρ, the so-called *virial expansion*:

$$p_{inter} = k_B T \sum_{i=2}^{\infty} B_i(T) \rho^i \tag{3.18}$$

where the $B_i(T)$ are called *virial coefficients*. For many purposes, it is sufficient to consider only the first term in this expansion, that is the second virial coefficient $B_2(T)$. We then have

$$F = N k_B T \left[\ln(\rho\lambda^3) - 1 + B_2\rho\right] \tag{3.19}$$

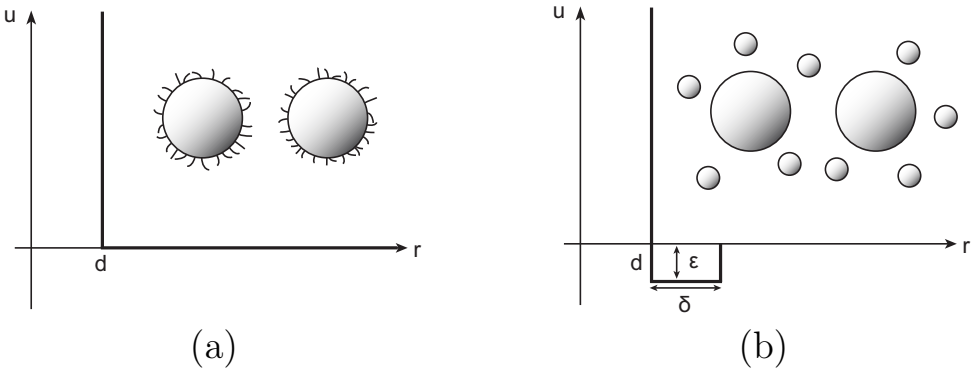
$$p = \rho k_B T \left[1 + B_2\rho\right] \tag{3.20}$$

Next we discuss how $B_2(T)$ depends on the interaction potential $U(\mathbf{r})$. For this purpose, we focus on the special case of isotropic and pairwise additive potentials $u(r)$, that is

$$U(\mathbf{r}) = \sum_{i<j} u(|\mathbf{r}_i - \mathbf{r}_j|) \tag{3.21}$$

This class of potentials contains many important cases:

- hard spheres (e.g. sterically stabilized latex beads)
- square well potential (e.g. depletion interaction)
- Lennard-Jones potential (real gas with van der Waals-attraction and Born repulsion)
- Coulomb potential for charged particles (with screening by counterions: Yukawa potential)



(a) (b)

Figure 3.1: Colloids can be used to tailor model potentials from statistical mechanics. (a) Sterically stabilized colloids result in a hard sphere fluid. (b) If in addition there is a depletion interaction, then a square-well potential with small interaction range results.

- DLVO-potential for colloidal stability (van der Waals-attraction and screened electrostatic repulsion)

It does not contain orientation-dependent interaction potentials like rods forming liquid crystals or electric dipoles forming chains. Given the assumption of an isotropic and pairwise additive potential, a general formula can be derived for the second virial coefficient $B_2(T)$ (given here without derivation):

$$B_2(T) = -\frac{1}{2}\int d\mathbf{r}\left(e^{-\beta u(r)} - 1\right) \tag{3.22}$$

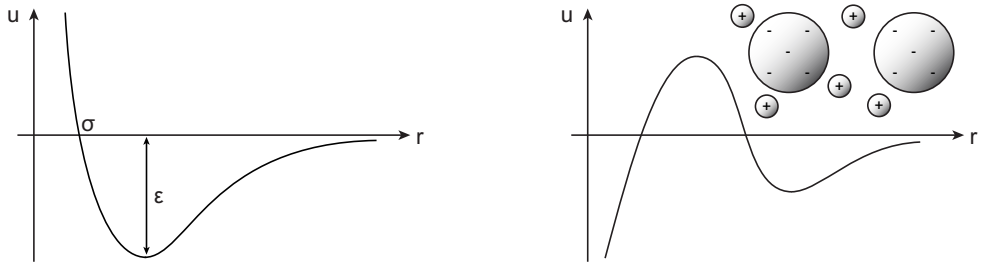
We now calculate $B_2(T)$ for several cases of interest.

3.2.1 Hard spheres

For hard spheres with diameter d, compare Fig. 3.1a, we get

$$B_2(T) = 2\pi\int_0^d drr^2 = \frac{2\pi}{3}d^3 = 4V_s > 0 \tag{3.23}$$

Thus B_2 has no temperature dependance, is four times the eigenvolume $V_s = \pi d/6$ of the spheres and is positive. Because then $p_{inter} = B_2\rho^2 k_B T > 0$, pressure increases due to the excluded volume effect. It is important to note that for hard systems, all thermodynamic effects are of entropic origin and there is no temperature dependance at all because there is no finite interaction energy.



(a) (b)

Figure 3.2: Smooth isotropic and pairwise interaction potentials common in colloidal physics. (a) The Lennard-Jones potential is the standard choice to model the van der Waals-interaction between molecules and results in a similar phase diagram as the square-well potential. (b) The DLVO-potentials describes the interaction between charged colloids surrounded by an electric double layer.

3.2.2 Attractive energy

Next we consider an attractive well of depth $-\epsilon$ located between d and $d+\delta$. Then

$$B_2(T) = -2\pi \int_d^{d+\delta} dr r^2 (e^{\beta\epsilon} - 1) \approx -2\pi (d^2\delta) \frac{\epsilon}{k_B T} \tag{3.24}$$

where the approximation is valid for $\delta \ll d$ and $\epsilon \ll k_B T$. We conclude that now the second virial coefficient is negative, that is pressure decreases due to the attraction. This effect increases with increasing interaction energy ϵ, decreasing temperature T or increasing interaction range δ.

3.2.3 Square well potential

We finally consider a square well potential with hard core repulsion for $r < d$. In practise, this might be realized by combining sterically stabilized colloids with a short-ranged depletion interaction resulting from the addition of polymers to the solution, compare Fig. 3.1b. Now $B_2(T)$ is simply the sum of the two preceding ones:

$$B_2(T) = \frac{2\pi}{3} d^3 - 2\pi (d^2\delta) \frac{\epsilon}{k_B T} = b - \frac{a}{k_B T} \tag{3.25}$$

where we have introduced two positive constants b (the repulsive eigenvolume) and a (representing the attractive part). This general form of $B_2(T)$ has been confirmed experimentally for many real gases. It now allows to rewrite the gas law in the following way:

$$pV = Nk_BT(1 + B_2\frac{N}{V}) \tag{3.26}$$

$$= Nk_BT(1 + b\frac{N}{V}) - \frac{N^2a}{V} \tag{3.27}$$

$$\approx \frac{Nk_BT}{1 - b\frac{N}{V}} - \frac{N^2a}{V} \tag{3.28}$$

thus

$$p = \frac{k_BT}{(v - b)} - \frac{a}{v^2} \tag{3.29}$$

where $v = V/N = 1/\rho$ is the volume per particle. This is the *van der Waals equation of state*: the volume per particle is reduced from v to $v - b$ due to excluded volume, and pressure is reduced by the attractive interaction, that is less momentum is transfered onto the walls due to the cohesive energy.

3.3 Fluid-fluid phase transition

The van der Waals equation of state Eq. (3.29) is characterized by an instability. For a stable system, if a fluctuation occurs to higher density (smaller volume), then a larger pressure should result, which can counteract the fluctuation. Therefore thermodynamic stability requires

$$\left(\frac{\partial p}{\partial V}\right)_{T,N} < 0 \tag{3.30}$$

This criterion is equivalent to the statement that the free energy has to be a convex function of volume, otherwise the system could gain free energy by splitting into two. However, below the critical temperature $T_c = (8a)/(27bk)$ the van der Waals isotherms from Eq. (3.29) indeed have sections in which this stability criterion is violated. This is shown in Fig. 3.3a. Therefore there is a region of phase space in which the system is unstable and starts to decompose into two different fluid phases, each of which are located in stable regions. The requirement of equal pressure and chemical potential at phase

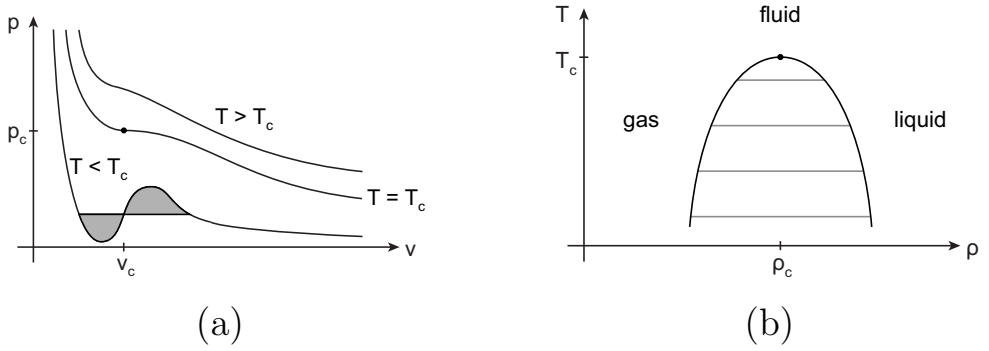


Figure 3.3: (a) The van der Waals equation of state contains an instability because for $T < T_c$, there are sections where $\partial p/\partial v$ is positive. (b) An equal area construction leads to a $T - \rho$ phase diagram with coexistence between gas and liquid phases at low temperature. The critical point marks a continuous (second order) phase transition.

coexistence leads to the equal area construction. Fig. 3.3b depicts the resulting phase diagram. At $T > T_c$, only one fluid phase exists. At the critical point with T_c, $v_c = 3b$ and $p_c = a/(27b^2)$, a second order phase transition takes place, that is the transition between the two fluids is continuous. At $T < T_c$, the transition is 1st order, that is there is a jump in v. The low and high density fluids correspond to gas and liquid, respectively.

3.4 Fluid-solid phase transition

The fluid-fluid phase transition is driven by the attractive part of the interaction potential. There is another phase transition described by the van der Waals equation of state which in contrast is driven by the repulsive part. If we consider the first term in the van der Waals equation of state Eq. (3.29), we can define a free volume

$$V - Nb = V(1 - \rho b) = \alpha_f V \tag{3.31}$$

where

$$\alpha_f = 1 - \frac{\rho}{\rho_0} \tag{3.32}$$

is the free volume fraction and $\rho_0 = 1/b = 3/(2\pi d^3)$ is the maximal density in the fluid phase. In order to estimate the free volume in the solid phase,

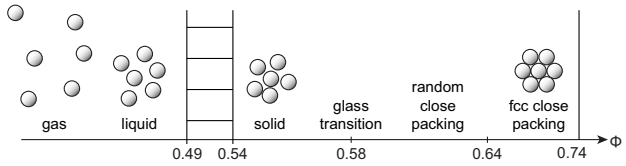


Figure 3.4: The phase diagram of hard spheres as obtained from computer simulations has no temperature-dependance and only depends on volume fraction. The liquid-solid phase phase transition is driven by entropy. The glass transition and the transition from random close packing to face centered cubic (fcc) close packing are structural rather than thermodynamic in nature.

we consider a simple cubic arrangement of spheres. The cube dimension is $V^{1/3}$ and it contains a smaller cube free of spheres with dimension $V^{1/3} - d$. Therefore

$$\alpha_s V \approx (V^{1/3} - d)^3 \Rightarrow \alpha_s = \left(1 - \left(\frac{\rho}{\rho_{cp}}\right)^{1/3}\right)^3 \tag{3.33}$$

where the maximal density in the solid phase can be chosen to be the close packing density $\rho_{cp} = 0.74(6/\pi d^3) \approx 3\rho_0$ (0.74 is the maximal volume fraction in the closest sphere packing, which according to Kepler's conjecture is the face centered cubic arrangement, that is the famous cannon ball or orange stacking). The different values of the maximal densities for fluid and solid phases combined with the different scaling with ρ leads to the result that the solid phase has a larger free volume at higher density. Because the free energy

$$F = Nk_BT\left[\ln\left(\frac{N\lambda^3}{\alpha V}\right) - 1\right] \tag{3.34}$$

directly follows from the free volume fraction α, there is a fluid-solid phase transition with increasing density. Somehow surprisingly, the solid phase is entropically more favorable at large densities because the ordering creates larger cavities of free volume than it is possible in a disordered fluid. This shows that entropy, although in general it is the tendency to increase disorder, can lead to local ordering. Because this effect does not require attractive energy, it does not depend on temperature. It is nicely confirmed for hard spheres by experiments on colloidal systems as well as by computer simulation, which both locate the fluid-solid transition at the packing

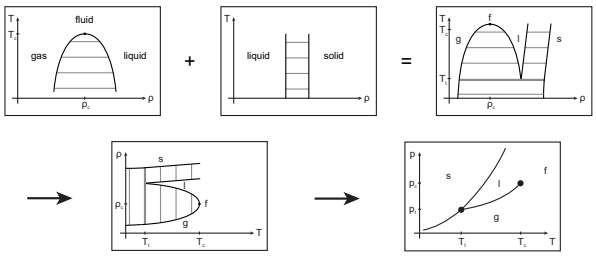


Figure 3.5: Typical phase diagram for a one-component system like a Lennard-Jones fluid. Fluid-fluid and fluid-solid phase transitions together result in the typical camel-shape phase diagram. The triple line at the triple temperature T_t denotes a three-phase coexistence. In $p-T$ space, it becomes a triple point.

fractions of 0.494 and 0.545. Note that there is no fluid-fluid transition for hard spheres because there is no attraction. The phase diagram from Fig. 3.4 shows that in addition to this thermodynamic phase transition, there are also two structural transitions (glass transition and the transition from random close packing to face centered cubic close packing) at higher densities.

3.5 Simple fluids

Since the square well potential leads to the van der Waals equation of state, both the fluid-fluid and the fluid-solid phase transitions are present in this case. The best studied model for a simple one component fluid is the Lennard-Jones fluid, compare Fig. 3.2a, which also has both excluded volume and attraction:

$$u(r) = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] \tag{3.35}$$

Here the attractive part represents the van der Waals interaction, while the repulsive part represents Born repulsion. The potential minimum $-\epsilon$ is reached at $r = 2^{1/6}\sigma \approx 1.12\sigma$. For Argon, one has $\epsilon = 0.4\ k_B T$ and $\sigma = 0.34$ nm. The big advantage of the Lennard-Jones fluid over the square well potential is its continuous nature, which makes it very well suited for computer simulations. In Fig. 3.5 we show the resulting phase diagram which is typical for a one-component fluid. There are two phase transitions, the gas-liquid phase transition driven by the attractive energy and the liquid-solid phase transition driven by entropy. Together they result in a triple line of gas-liquid-solid phase coexistence at the triple temperature T_t. If the

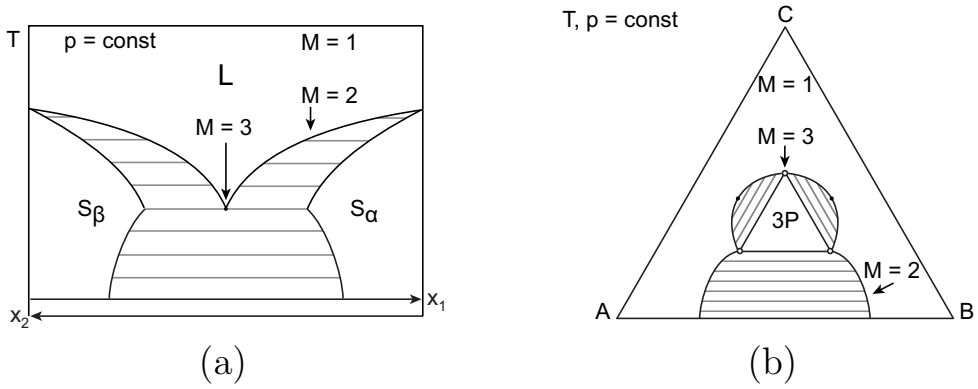


Figure 3.6: (a) Typical phase diagram for a two-component system at constant pressure. The eutectic point has three-phase coexistence (M=3) between two solids and a liquid phase. The lines represent two-phase coexistence (M=2). At high temperature, only one phase exists (M=1). (b) Typical phase diagram of a three-component system at constant temperature and pressure. The Gibbs triangle describes the concentrations of the three components A, B and C, which represent for example water, oil and amphiphile. Here we assume A-B symmetry for simplicity. At low concentration of C, A and B demix in a two-phase coexistence region (delineated by the M=2 lines), but at high concentration of C, they form a single phase (M=1). At intermediate concentrations of C, there is three phase coexistence (3P), with the three points of the triangle corresponding to three phase points (M=3).

typical camel-shape phase diagram is converted into $p-T$-space, the triple line becomes a triple point.

3.6 Multiple component fluids

Phase diagrams of multicomponent systems can be very complicated, but they are very common in physical chemistry and biological systems. For r components and M phases, the Gibbs phase rule states that $f = 2 - M + r$ variables can be assigned freely. For a one compoment system, $r = 1$ and $f = 3 - M$. Therefore coexistence of up to three phases is possible. Indeed the phase diagram for the Lennard-Jones fluid leads to a triple point at which gas, fluid and solid coexist. For a two component system, $r = 2$ and $f = 4 - M$, that is now up to four phases are possible. Fig. 3.6a shows a typical phase diagram for constant pressure. Then up to three-

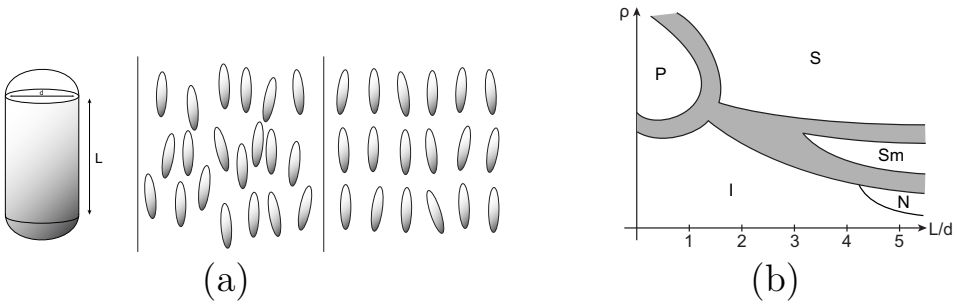


Figure 3.7: (a) Hard spherocylinders are a model system for liquid crystals and can form nematic and smetic phases. In a nematic liquid crystal, all molecules tend to align in the same direction. In a smectic liquid crystal, in addition they order into layers. (b) Phase diagram of hard spherocylinders as a function of aspect ratio L/d. Two-phase regions are shaded. I isotropic, S solid, P plastic solid, Sm smectic and N nematic.

phase coexistence is possible, which is realized here at the eutectic point (M=3). For a three-component system, $r = 3$ and $f = 5 - M$. Fig. 3.6b shows for a mixture symmetric in A and B that for constant temperature and pressure, three-phase coexistence (3P) can occur. An important example of such a system is a mixture of water (A), oil (B) and amphiphile (C). The phase diagram shows that the amphiphile can induce water and oil to mix if present at sufficiently high concentrations.

3.7 Liquid crystals

Other soft matter systems with well studied phase behaviour are lipid-water mixtures, diblock copolymers and liquid crystals. In these kind of systems, usually the components do not have simple spherical shapes. For example, lipids and diblock copolymers are both elongated molecules with two different parts (e.g. water-like and oil-like parts for lipids), leading to amphiphilic phase behaviour characterized by a large amount of internal interface. The tendency to spontaneously form internal interfaces is also typical for liquid crystals, which already arise for relatively simple sytems like hard spherocylinders, compare Fig. 3.7. The main classes of liquid crystal phases are nematic and smectic phases. In nematic phases, rotational symmetry is broken and all molecules align in average in the same direction. In smectic

phases, in addition positional ordering occurs, resulting in stacks of aligned molecules. The transition between the isotropic and nematic phases is second order and has been described first by Onsager's analytical theory. Again it is pure entropy which leads to an ordering phenomena.

Chapter 4

Interfaces

In finite sized systems, two coexisting phases have to develop some interface between them. In soft matter physics, often the physics of the bulk phases is sufficiently simple that the system properties are determined by its interfaces. In this chapter, we will introduce the concepts required to deal with this situation. We start on the mesoscopic level with a Ginzburg-Landau model and then take the limit to the effective interface model. We introduce the basic tools from differential geometry required to deal with the geometry of surfaces and then discuss the main features of surfaces under tension, including thermally activated capillary waves on a flat surface and the Rayleigh-Plateau instability of cylindrical surfaces. We also discuss a few other tension-related issues, including soap bubbles, foams, wetting and cell shapes.

4.1 ϕ^4-model

Interfaces can be modeled on different length scales. On the scale of single particles, one can either use off-lattice models like Monte Carlo simulations of Lennard-Jones fluids or lattice models like the binary gas (Ising) model. On a mesoscopic level one can use locally averaged densities. Since these densities vary in space, here one essentially deals with field theories. In another coarse graining step, one arrives at effective interface models, where the interface is described approximated by a mathematical surface. Fig. 4.1 schematically depicts these different scales.

We now start at the mesoscopic scale and at the vicinity of a critical

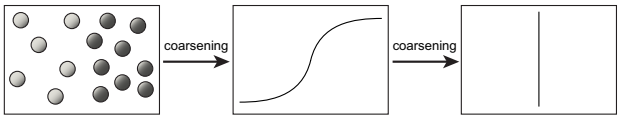


Figure 4.1: An interface between two different phases (like gas and liquid) can be modeled on different scales. From left to right: particle level - level of density fields - effective interface level.

point for a fluid-fluid phase coexistence. We assume that there exists an order parameter ϕ which has different values in the two fluid phases below T_c. Above T_c, only one phase exists and the order parameter can only obtain one value, which we take to be 0. The simplest model for this szenario is the free energy density

$$f(\phi) = \frac{r}{2}\Phi^2 + s\Phi^4 \tag{4.1}$$

If $r \sim (T - T_c)$, then for $T < T_c$ a Mexican hat shape arises and two minima coexist. Away from the phase boundary, one or the other minimum dominates. For $T > T_c$, we have a simple parabola and $\phi = 0$ is the only minimum. Exactly at $T = T_c$, f becomes very flat and large fluctuations arise. In Fig. 4.2, we show how the different potential shapes correspond to different points in the phase diagram of a one-component fluid. In this case, the order parameter Φ can be identified with the density different between gas and liquid phase.

If interfaces exist, then we have to add a penalty to the free energy for having an interface. Because an interface requires a spatially varying profile, the free energy now has to be a functional of $\phi(\mathbf{r})$. The simplest assumption possible is a gradient-squared term:

$$F[\phi] = \int d\mathbf{r} \ \left\{ \frac{c}{2}(\nabla\phi)^2 + f(\phi) \right\} \tag{4.2}$$

This ϕ^4-*model* is the best studied example of a Ginzburg-Landau theory and can be derived from the Ising model (note the symmetry under $\phi \to -\phi$).

We now consider one interface which extends in the z-direction only. Then $\phi(z)$ varies from one minimum $\phi_0 = (-r/4s)^{1/2}$ at one side to the other minimum $-\phi_0$ at the other side. We define the energy per interfacial area, the so-called *surface tension*, as

$$\sigma = \frac{F}{A} = \int dz \left\{ \frac{c}{2}\phi'^2 + \tilde{f}(\phi) \right\} \tag{4.3}$$

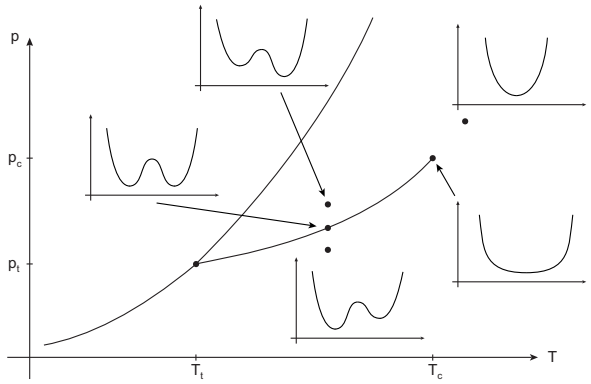


Figure 4.2: The free energy density f as a function of order parameter Φ for different regions of the on-component phase diagram.

where $\tilde{f} = f - f_0$ with $f_0 = -r^2/16s$ is the free energy density measured relative to the ground state.

We now calculate σ as a function of the parameters of the ϕ^4-model. Before we demonstrate the exact solution, we start with a scaling approach. The energy barrier between the two minima scales as $\epsilon \sim r^2/s$. The distance between the two minima scales as $\Delta\Phi \sim (-r/s)^{1/2}$ (compare Fig. 4.3a). With an interfacial width l, the surface tension then scales as

$$\sigma \sim c\left(\frac{\Delta\Phi}{l}\right)^2 l + \epsilon l \tag{4.4}$$

The optimal interfacial width l_0 is finite and results from the balance between the gradient term (wants it large) and the energy penalty for not being in the minimum (wants it small):

$$l_0 \sim \left(\frac{c}{-r}\right)^{1/2} \tag{4.5}$$

By reinserting this into the free energy expression, we find the surface tension as

$$\sigma \sim \left(\frac{cr^3}{s^2}\right)^{1/2} \tag{4.6}$$

Note that surface tension σ vanishes at the critical point $r = 0$, as it should.

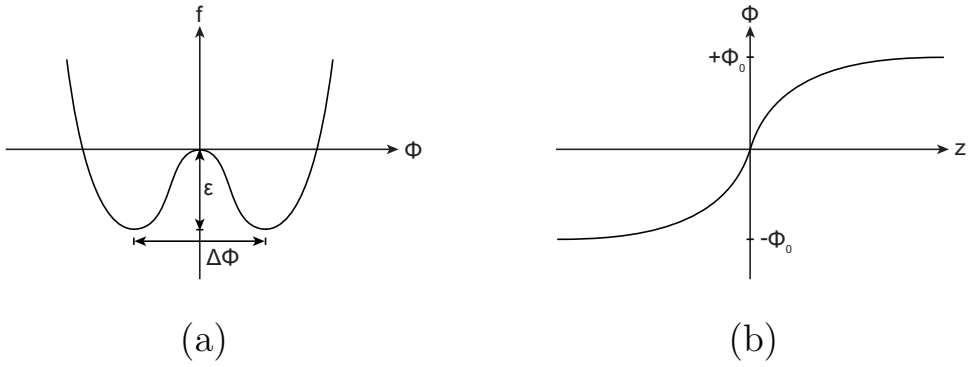


Figure 4.3: (a) Free energy density f as a function of the order parameter Φ for the Φ^4-model at phase coexistence. (b) Kink solution of the Φ^4-model for the one-dimensional interface.

Next we turn to the exact solution. The Euler-Lagrange equation for the ϕ^4-model reads

$$\frac{\delta F[\phi]}{\delta \phi} = -c\phi'' + \tilde{f}'(\phi) = 0 \tag{4.7}$$

Multiplying with ϕ' we see that there is an integral of motion:

$$\phi' - \frac{2\tilde{f}}{c} = B \tag{4.8}$$

where B is a constant of integration. Another integration gives

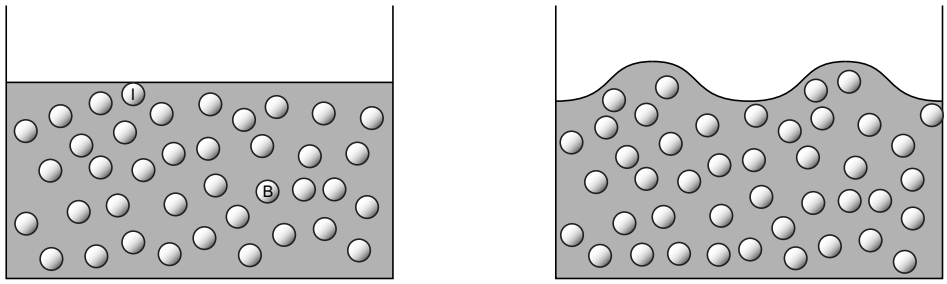
$$z - z_0 = \int_{\phi(z_0)}^{\phi(z)} \frac{d\phi}{(\frac{2\tilde{f}}{c} + B)^{1/2}} \tag{4.9}$$

As boundary conditions we have $\phi(\pm\infty) = \pm\phi_0$ and $\phi'(\pm\infty) = 0$. The latter conditions leads to $B = 0$. Moreover we choose $z_0 = 0$ and $\phi(z_0) = 0$. Noting that one can write

$$\tilde{f} = \frac{-r}{4}\frac{(\phi^2 - \phi_0^2)^2}{\phi_0^2} \tag{4.10}$$

we now can solve the integral:

$$z = \left(\frac{2c}{-r}\right)^{1/2} \int \frac{\phi_0 d\phi}{\phi^2 - \phi_0^2} = \left(\frac{c}{-2r}\right)^{1/2} \tanh^{-1}\frac{\phi}{\phi_0} \tag{4.11}$$



(a) (b)

Figure 4.4: (a) Physical origin of surface tension: particles at the interface (I) have lower coordination than in the bulk (B). (b) Thermal energy leads to interface fluctuations (capillary waves).

Inversion gives the famous *kink solution* for the interfacial profile:

$$\phi(z) = \phi_0 \tanh \frac{z}{l_0} \tag{4.12}$$

with $l_0 = (c/-2r)^{1/2}$. A schematic plot is shown in Fig. 4.3b. Inserting the kink solution into F gives an exact expression for the surface tension:

$$\sigma = \left(\frac{-cr^3}{18s^2}\right)^{1/2} \tag{4.13}$$

Thus the exact solution nicely corraborates the scaling analysis.

In general, the physical origin of surface tension is that particles at the interface have fewer favorable contacts with like particles than in the bulk, compare Fig. 4.4a. Therefore the surface tension σ for a simple fluid can be estimated to be the energy E per particle-particle interaction divided by particle size a squared. For complex fluids, we have to replace this by typical energy and length scales:

- simple fluid (eg water-oil): $E = kT$, $a = 3A$ (molecular) $\rightarrow$ $\sigma = 40\ erg/cm^2$
- complex fluid (eg microemulsion): $E = kT$, $a = 100A$ (mesoscopic) $\rightarrow \sigma = 0.04\ erg/cm^2$

Here $erg/cm^2 = mJ/m^2$ is a convenient unit for surface tension. Sensitivity to size comes from quadratic dependance on the characteristic length scale a. Earlier we have used a similar line of reasoning to argue that soft matter is characterized by small energy scales and large length scales, because elastic moduli scales as E/a^3. Experimentally one finds that the water-air interface has a surface tension of $\sigma = 72\ erg/cm^2$.

Usually σ is positive, otherwise the two bulk phases defining the interface would mix, because $\sigma < 0$ corresponds to a proliferation of interfaces. This situation can occur in systems with surfactants, if the chemical potential for surfactants at the interface is sufficiently favorable. Then an interface instability called *spontaneous emulsification* occurs, which usually is counteracted by the interfacial curvature energy. In order to understand why surfactants lower the surface tension, we consider a molecule which is attracted to the interface and then behaves like an ideal gas in the plane of the interface:

$$F = A(\sigma_0 + kT\rho(\ln(\rho a_0) - 1)) \tag{4.14}$$

where $\rho = N/A$ is surfactant area density, N is number of surfactant molecules, and a_0 is some molecular area. The effective surface tension then follows as

$$\sigma = \left(\frac{\partial F}{\partial A}\right)_{N=const} = \sigma_0 - kT\rho \tag{4.15}$$

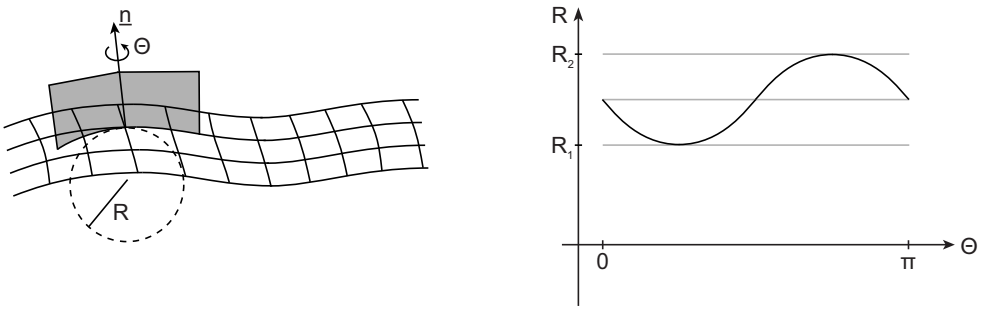
that is the entropy of the ideal gas lowers the surface tension. The same effect follows from excluded volume and repulsive interactions like charged headgroup repulsion.

4.2 Differential geometry of surfaces

The Hamiltonian of the effective interface model looks deceptively simple:

$$H = \sigma \int dA \tag{4.16}$$

However, the partition sum is the path integral over all possible configurations of the interface. In general, these kinds of calcuations are not simple. In order to proceed, one now has to deal with the geometrical properties of mathematical surfaces. To this end, we have to introduce a few concepts from differential geometry.



(a) (b)

Figure 4.5: Differential geometry of surfaces. (a) At any point on the surface, one can construct a normal vector n. Any plane containing this normal vector defines a direction θ on the surface, for which one can determine the radius of curvature R. (b) R as a function of θ has a minimum R_1 and a maximum R_2. Mean curvature is defined as $H = (1/R_1 + 1/R_2)/2$ and Gaussian curvature as $K = 1/(R_1 R_2)$, respectively.

A two-dimensional surface in three-dimensional space is defined by the position vector $\mathbf{f}(x, y) = (f_1(x, y), f_2(x, y), f_3(x, y))$, where x and y are the internal coordinates. The two tangential vectors $\partial_x \mathbf{f}$ and $\partial_y \mathbf{f}$ span the tangential plane, while the normal vector $\mathbf{n} = (\partial_x \mathbf{f} \times \partial_y \mathbf{f})/|\partial_x \mathbf{f} \times \partial_y \mathbf{f}|$ points perpendicularly to it. It we rotate a plane containing $\mathbf{n}$ around it, compare Fig. 4.5a, for each angle of rotation θ a curve is cut out of the surface, to which we can fit a circle with radius $R(\theta)$. The curvature is defined as $\kappa(\theta) = 1/R(\theta)$. Since $\kappa(\theta)$ is a smooth function on a compact carrier, it must have a minimum and a maximum, compare Fig. 4.5b. This defines the principal curvatures κ_1 and κ_2 at the given point on the surface. They correspond to the two principal radii of curvature, $R_1 = 1/\kappa_1$ and $R_2 = 1/\kappa_2$. The *mean curvature* H and the *Gaussian curvature* K are defined as

$$H = \frac{\kappa_1 + \kappa_2}{2} = \frac{1}{2}\left(\frac{1}{R_1} + \frac{1}{R_2}\right), \qquad K = \kappa_1 \kappa_2 = \frac{1}{R_1 R_2} \tag{4.17}$$

The mean curvature H describes how strongly the surface bends on average. At least equally important, the Gaussian curvature K denotes the kind of geometry one is dealing with, as shown in Fig. 4.6:

- elliptic geometry represented by a sphere: $R_1 = R_2 = R$, $H = -1/R$, $K = 1/R^2 > 0$

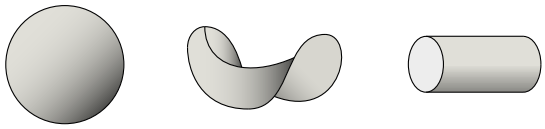


Figure 4.6: Three fundamentally different kinds of geometries can be distinguished according to the Gaussian curvature K. (a) A sphere has $K > 0$. (b) A saddle has $K < 0$. (c) A cylinder has $K = 0$.

- hyperbolic geometry represented by a saddle: $R_1 = -R_2$, $H = 0$, $K = -1/R_1^2 < 0$
- parabolic geometry represented by a cylinder: $R_2 = \infty$, $H = -1/2R$, $K = 0$

Note that the two minus signs for H are conventions, because we define curvature to be negative if the surface bends away from the normal, which we take to point outwards. The *theorema egregium* by Gauss states that K depends only on the inner geometry of the surface, that is it does not depend on the definition of a normal vector pointing in the surrounding space. The *Gauss-Bonnet theorem* states that K integrated over a closed surface is a topological constant:

$$\int dAK = 2\pi\chi \tag{4.18}$$

where $\chi = 2(1 - G)$ and G are integer numbers called *Euler characteristic* and *genus*, respectively. The genus denotes the number of holes in the object. So for the sphere, we have $G = 0$ and $\chi = 2$, and indeed Eq. (4.18) specified to a sphere gives $(4\pi R^2)(1/R^2) = (2\pi)2$.

In order to evaluate integrals like the one in Eq. (4.18) for arbitrary surfaces, we need formulae for the area element dA, the mean curvature H and the Gaussian curvature K as a function of the internal coordinates x and y. For this purpose, we introduce three 2x2-matrices:

$$g_{ij} = \partial_i \mathbf{f} \cdot \partial_j \mathbf{f},\ h_{ij} = -\partial_i \mathbf{n} \cdot \partial_j \mathbf{f},\ a = hg^{-1} \tag{4.19}$$

g is called the *first fundamental form* or the *metric tensor*, h is the *second fundamental form* and a is the *Weingarten matrix*. The area element follows from the metric tensor:

$$dA = |\partial_x \mathbf{f} \times \partial_y \mathbf{f}|\ dxdy = (\det g)^{1/2}\ dxdy \tag{4.20}$$

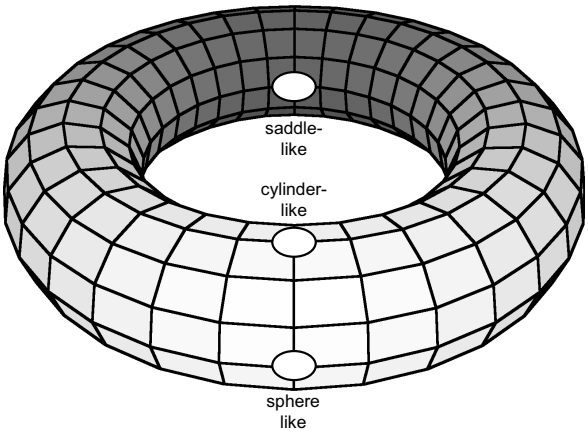


Figure 4.7: A torus has saddle, cylinder and sphere-like parts, depending on the internal coordinates.

and the curvatures follow from the Weingarten matrix:

$$H = \frac{1}{2}\ tr\ a \qquad\qquad K = \det\ a = \frac{\det h}{\det g} \tag{4.21}$$

We now can apply these concepts for any geometry of interest. For the cylinder, we have $x = \varphi$ and $y = z$ with $\mathbf{f} = (R\cos\varphi, R\sin\varphi, z)$. This leads to $dA = Rd\varphi dz$, $H = -1/2R$ and $K = 0$. For the sphere, we have $x = \varphi$ and $y = \theta$ with $\mathbf{f} = R(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$. This leads to $dA = R^2\sin\theta d\varphi d\theta$, $H = -1/R$ and $K = 1/R^2$. In these two examples, the curvatures are constant over the surface. Therefore cylinder and sphere are examples for *surfaces of constant mean curvature* (CMS-surfaces). There is a theorem which states that the sphere is the only CMC-surface which is compact.

As an example for a geometry with varying curvature, we consider the torus, for which we have $x = \varphi$ and $y = \theta$ with $\mathbf{f} = (t\cos\varphi, t\sin\varphi, r\sin\theta)$, where $t = a + r\cos\theta$. Here r and $a > r$ are the outer and inner radii of the torus, respectively. This leads to

$$g = \begin{pmatrix} t^2 & 0 \\ 0 & r^2 \end{pmatrix},\ h = \begin{pmatrix} -t\cos\theta & 0 \\ 0 & -r \end{pmatrix},\ a = \begin{pmatrix} -\cos\theta/t & 0 \\ 0 & -1/r \end{pmatrix} \tag{4.22}$$

from which we get

$$dA = trd\varphi d\theta,\ H = -\frac{1}{2}\left(\frac{\cos\theta}{t} + \frac{1}{r}\right),\ K = \frac{\cos\theta}{tr}\ . \tag{4.23}$$

Because the torus is invariant under a rotation with φ, the results for the curvatures only depend on θ. The Gaussian curvature now varies over the surface. For $\theta = 0$ we are on the outer side, which is sphere-like with $K = 1/(ar + r^2) > 0$. For $\theta = \pi/2$ and $\theta = 3\pi/2$ we are on the top and bottom sides, which are cylinder-like with $K = 0$. And for $\theta = \pi$ we are on the inner side, which is saddle-like with $K = -1/(ar - r^2) < 0$. Overall area is $\int dA = 4\pi^2 ar$ and $\int dAK = 0$, in accordance with the Gauss-Bonnet theorem for $\chi = 0$ ($G = 1$).

An important geometry for the following is the *Monge parametrization* for a surface which can be described by a height function $h(x, y)$ (that is there should be no overhangs). Starting with $\mathbf{f} = (x, y, h(x, y))$, one can derive exact, but complicated formulae for dA, H and K. In the approximation of a nearly flat surface ($h_x, h_y \ll 1$), they simplify to

$$dA \approx \sqrt{1 + h_x^2 + h_y^2}\; dxdy \;\approx [1 + \frac{1}{2}(h_x^2 + h_y^2)]\; dxdy \tag{4.24}$$

$$H \approx \frac{1}{2}(h_{xx} + h_{yy}),\;\; K \approx h_{xx}\; h_{yy} - (h_{xy})^2 \tag{4.25}$$

Finally it is instructive to consider small variations around a given surface:

$$\mathbf{f}^t = \mathbf{f} + t\varphi\mathbf{n} \tag{4.26}$$

where t is a small parameter and $\varphi(x, y)$ an arbitrary scalar field. Using the concepts introduced above, we can calculate the resulting change in the area element to first order in t:

$$dA^t - dA = -2t\varphi(x, y)H(x, y)(\det g)^{1/2}dxdy \tag{4.27}$$

For this first variation to vanish, we must have $H = (\kappa_1 + \kappa_2)/2 = 0$ everywhere. A surface with vanishing mean curvature H is called a *minimal surface*, because stationarity under surface variation is a necessary (but not sufficient) condition for minimal surface area. From $H = 0$ we get $\kappa_1 = -\kappa_2$ and $K = -\kappa_1^2 < 0$, therefore minimal surfaces are saddle-like everywhere. From this property it follows directly that minimal surfaces without boundaries cannot be compact, because the saddle-like shape would cut any container in which one tries to enclose the surface. Examples for non-periodic minimal surfaces are the plane and the catenoid. Most minimal surfaces known however are periodic (including the singly-periodic helicoid, the doubly-periodic

Scherk-surface and the triply-periodic gyroid). Another important consequence of Eq. (4.27) is that surface and volume are related by curvature: because

$$dV^t - dV = t\varphi(x,y)(\det g)^{1/2}dxdy \tag{4.28}$$

we have $dA/dV = -2H$.

4.3 Capillary waves

We now consider the effective interface model in the Monge parametrization for nearly flat surfaces, compare Fig. 4.4b. Using Eq. (4.24), we have

$$F = \sigma \int dA = \sigma A_0 + \frac{1}{2}\sigma \int dxdy\ (h_x^2 + h_y^2) \tag{4.29}$$

Since the second term is always positive, the interface is stable under capillary waves. However, thermal fluctations will always be present, due to the gain in entropy. We first change to Fourier space:

$$\Delta F = \frac{1}{2}\sigma \int d\mathbf{q}\ q^2 |h(\mathbf{q})|^2 \tag{4.30}$$

Since the different q-modes decouple, the partition function can be calculated exactly (Gauss integral). The result is an example of the equipartition theorem (each harmonic mode has the energie $kT/2$):

$$\langle |h(\mathbf{q})|^2 \rangle = \frac{kT}{\sigma q^2} \tag{4.31}$$

We can now transform back to real space:

$$\langle h(\mathbf{r})^2 \rangle = \frac{kT}{\sigma}\frac{1}{(2\pi)^2}\int_{2\pi/L}^{2\pi/a} d\mathbf{q}\frac{1}{q^2} = \frac{kT}{2\pi\sigma}\ln\frac{L}{a} \tag{4.32}$$

Note that we deal with a two-dimensional system and that $\langle h(\mathbf{r})^2 \rangle$ does not depend on position $\mathbf{r}$ due to translational invariance. Also we had to introduce two cutoffs, a molecular length a and the system size L. The interface is rough for all temperatures, since its mean thickness increases logarithmically with system size. Yet, this divergence is relatively weak: for $\sigma = 100\ erg/cm^2$, $a = 3\ A$, $L = 100\ A$, the rms fluctuation is just 1.5 A, and for $L = 1\ cm$, it increases to 7.5 A only.

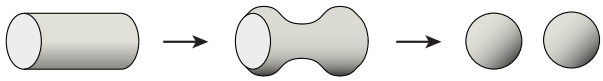


Figure 4.8: Rayleigh-Plateau instability: a cylinder under tension is unstable under long wavelength modulations and breaks up into spherical droplets.

We now consider the effect of gravity on an interface under tension: the gravitational energy per unit area is $\int_0^h dz\ zg\rho = g\rho h^2/2$. Therefore the free energy now reads

$$\Delta F = \frac{1}{2}\sigma \int dxdy \left\{ (h_x^2 + h_y^2) + (\frac{h}{\xi})^2 \right\} \tag{4.33}$$

where $\xi = (\sigma/\rho g)^{1/2}$ is the capillary length. For water, it is of the order of millimeter. The problem is still Gaussian, and we have

$$\langle |h(\mathbf{q})|^2 \rangle = \frac{kT}{\sigma(q^2 + \xi^{-2})} \tag{4.34}$$

Doing the inverse transform, we see that the logarithmic divergence is now cut off by the capillary length, like before by the system size, $\langle h(\mathbf{r})^2 \rangle \sim \ln(\xi/a)$.

4.4 Rayleigh-Plateau instability

A cylinder is unstable under long-wavelength capillary waves and breaks up into droplets, which have a lower surface/volume ratio. Consider a cylinder oriented in z-direction with a small undulation added:

$$\mathbf{r} = \begin{pmatrix} r(z)\cos\phi \\ r(z)\sin\phi \\ z \end{pmatrix} \tag{4.35}$$

with

$$r(z) = R(1 + \epsilon \sin(kz)) \tag{4.36}$$

where $k = 2\pi/L$ is the wavelength of the deformation. The unit cell volume of this cylinder is

$$V = \pi \int_0^L dz\ r(z)^2 \tag{4.37}$$

$$= \pi L \int_0^1 dz'\ R^2(1 + \epsilon^2 \sin^2(2\pi z')) \tag{4.38}$$

$$= \pi L R^2(1 + \frac{\epsilon^2}{2}) \tag{4.39}$$

where a term linear in ϵ dropped out. If the initial radius is R_0 and volume is conserved, then

$$R_0^2 = R^2(1 + \frac{\epsilon^2}{2}) \tag{4.40}$$

that is, since $\epsilon \ll 1$,

$$R = R_0(1 - \frac{\epsilon^2}{4}) \tag{4.41}$$

It is easy to calculate the metric tensor:

$$g_{ij} = \begin{pmatrix} r^2 & 0 \\ 0 & 1 + r_z^2 \end{pmatrix} \tag{4.42}$$

Therefore the surface energy is

$$F = \sigma \int d\phi dz \sqrt{detg} = 2\pi\sigma \int_0^L dz r \sqrt{1 + r_z^2} \tag{4.43}$$

$$= 2\pi\sigma L \int_0^1 dz' R(1 + \frac{R^2\epsilon^2}{2L^2} \cos^2(2\pi z')) \tag{4.44}$$

$$= 2\pi\sigma L R(1 + \frac{R^2\epsilon^2}{4L^2}) \tag{4.45}$$

where again a term linear in ϵ dropped out and we expanded the square root. The change in surface energy thus is

$$\Delta F = 2\pi\sigma L R(1 + \frac{R^2\epsilon^2}{4L^2}) - 2\pi\sigma L R_0 \tag{4.46}$$

$$= 2\pi\sigma L R_0 \frac{\epsilon^2}{4}(\frac{R_0^2}{L^2} - 1) \tag{4.47}$$

where we have used Eq. (4.41) and dropped higher terms in ϵ. We conclude that ΔF becomes negative (that is the systems becomes unstable) for long wavelength modulations, $L > R_0$.

(a) (b)

Figure 4.9: Soap bubble versus vesicle. (a) Soap films are surfaces under tension. Although the soap bubble is a closed object, area is not conserved due to variable film thickness. The bubble is stabilized against collapse by internal pressure. (b) Lipid bilayers are surfaces governed by bending energy. For a vesicle, area is conserved because bilayer thickness is constant. In practise, the presence of dirt leads to an effectively conserved volume.

4.5 Tension-related phenomena

The paradigm for a tension-dominated interface is the water-air interface. A similar system are soap films, that is thin films of water where the two water-air interfaces are stabilized by soap molecules. The overall system again acts as a surface under tension. In a soap bubble, interfacial tension is counterbalanced by increased pressure inside, therefore $\sigma dA = pdV$. From a thermodynamic point of view, this corresponds to minimization of the overall energy, that is $dF = \sigma dA - pdV = 0$. If radius R is changed by dR, area changes by $dA = 8\pi R dR$ and volume by $dV = 4\pi R^2 dR$. Thus we arrive at the *Laplace law* for spherical objects:

$$p = \frac{2\sigma}{R} \tag{4.48}$$

For arbitrary surfaces with curvatures varying over the surface, we can use the general relation between variations in area and volume resulting from Eq. (4.27). This leads to the general Laplace law:

$$\frac{dA}{dV} = \frac{p}{\sigma} = -2H \Rightarrow p = -2\sigma H \tag{4.49}$$

For the sphere, $H = -1/R$ and we recover Eq. (4.48). In general, Eq. (4.49) predicts that soap films between regions with different pressures are CMC-

surfaces. In mathematical terms, this corresponds to the statement that surfaces with minimal surface area and under a volume constraint are CMC-surfaces. Because the sphere is the only compact CMC-surface, liquid droplets and soap bubbles are spherical. For soap films spanning open frames (like a soap film connecting two opposing wire rings), the pressure difference p vanishes, $H = 0$ and we deal with a minimal surface (like the catenoid for the two rings).

Dry foams are ensembles of soap films meeting along edges, with the edges meeting in vertices (*wet foams* are spherical gas bubbles in an aqueous matrix). In contrast to the situation with soap bubbles, which are forced to have Laplace pressure, here pressure can equilibrate with time by gas diffusion over the water films. Therefore the faces are minimal surfaces. In many cases, they are rather flat, leading to polyhedral arrangements. Foams obey some fundamental rules which have been stated first by Plateau and later have been proved rigorously:

- always three faces meet in an edge, with angles of 120°
- always four edges (that is six faces) meet in a vertex, with the tetrahedral angle of 109°

Today foams can be made from nearly any kind of material (liquid, polymers, metal, ceramics, glass, etc), but usually one sees the signature of these rules, since the synthesis often is based on some liquid-evolving process.

Two-dimensional foams (between glass plates) lead to the honeycomb structure, because it obeys the first rule by Plateau. In three dimensions, the optimal solution, that is the equal-volume partitioning of space with minimal surface area, is elusive (*Kelvin problem*). These are the main candidates discussed until today:

- regular dodecahedron (12 pentagons): does not fill space
- rhombic dodecahedron (Wigner-Seitz cell for the face centered cubic crystal): 12 faces meet at every vertex
- tetrakaidecahedron (Wigner-Seitz cell for body centered cubic crystal, truncated octahedron): suggested by Kelvin in 1887
- Weaire-Phelan structure (six 14-sided Goldberg cells and 2 pentagonal dodecahedra): was shown in 1994 to be better by few percents than the Kelvin-foam

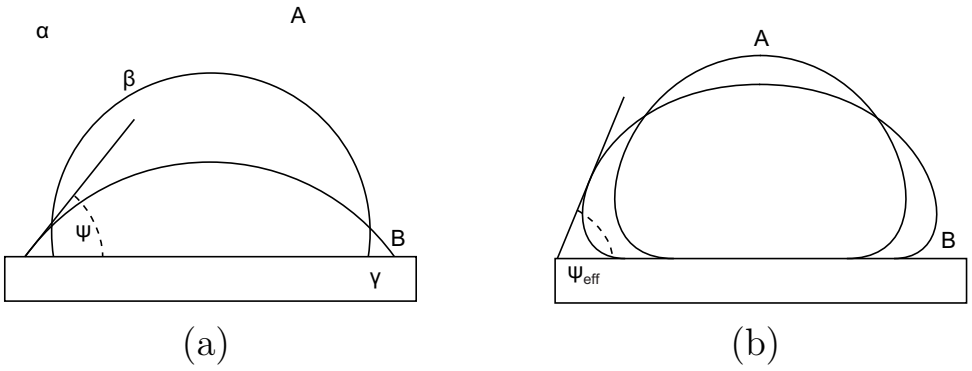


Figure 4.10: (a) A liquid droplet on a solid substrate has the shape of a spherical cap with a contact angle determined by the Young law. If the contact angle is larger than 90°, the liquid does not like the surface and we have dewetting (A). Otherwise the liquid likes the surface and we have wetting (B). (b) Because sharp bends are energetically unfavorable, a vesicle adhering to a solid substrate does not have a well-defined contact angle (A), except if the contact energy becomes very large (B).

In practise, it is very difficult to make three-dimensional monodisperse foams and real foams always have a statistics dominated by pentagonal faces.

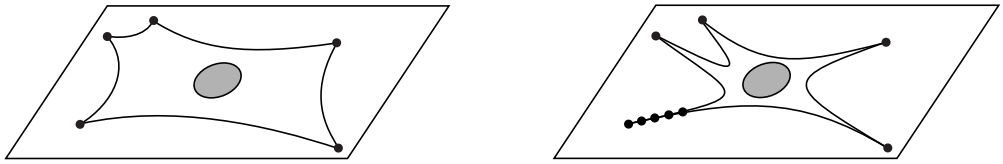
The interaction of liquid droplets with a substrate is called *wetting*, compare Fig. 4.10a. A wetting situation involves three phases: the liquid in coexistence with its vapor and a solid support. If the interaction between liquid and solid is sufficiently favorable, complete wetting occurs and a thin film forms on the surface. In the opposite case, dewetting occurs and a droplet rounds up into a sphere. In the intermediate case of partial wetting, a finite contact angle θ results. Since the system is tension-dominated, the Laplace law still holds true at any free point of the free surface. Therefore drops wetting a substrate are spherical caps. The contact angle is determined by the boundary conditions. Force balance dictates

$$\sigma_{SL} + \sigma_{VL} \cos\theta = \sigma_{SV} \tag{4.50}$$

leading to the *Young equation*:

$$\cos\theta = \frac{\sigma_{SV} - \sigma_{SL}}{\sigma_{VL}} \tag{4.51}$$

where θ is the *contact angle*. Therefore increasing σ_{VL} decreases θ and thus favors complete wetting. In practise, this is often counterbalanced by the



(a) (b)

Figure 4.11: Effect of tension in cell adhesion. (a) The shape of cells adhering to a flat and rigid substrate is often characterized by a sequence of inward-curved circular arcs resulting from a Laplace law modified for two dimensions. (b) When the actin cytoskeleton is weakened by specific drugs, the cell envelope retracts and the processes connecting the cell to the points of adhesion become cylinder-like. Then a Rayleigh-Plateau instability can occur.

effect of the van der Waals interaction, which favors thickening of the film. Effectively this leads to pancake-like droplet shapes, as does gravity.

The science of wetting is very advanced both in theoretical and experimental terms and is also of very large technological importance. Important subjects here are eg effects of van der Waals interactions and gravitation, contact angle hysteresis, precursor films, line tension effects, spreading dynamics and wetting on topographically or chemically structured surfaces. Other related subjects are *capillary forces*, which make the sap rise in trees, and *capillary bridges*, which prevent sandcastles from falling.

One important application of the concept of tension are the shapes of biological cells. Cell shapes often appear to be dominated by surface tension: cells tends to be round in solution, cell adhesion to a two-dimensional surface often feature circular arcs connecting points of adhesion (two-dimensional version of the Laplace equation), tissue cells often are polyhedral with a dominance of pentagonal faces (three-dimensional foam) and often tubular extensions of cells show a pearling instability (Rayleigh-Plateau instability). The origin of the tension in the cell envelope is not clear. Most likely it is a combination of several effects, including passive prestrain and active contractility by myosin molecular motors in the actin cytoskeleton, tension in the plasma membrane (possibly related to membrane reservoirs connected

to the plasma membrane), and an effective tension arising from the interaction between actin cortex and plasma membrane. In Fig. 4.11 we show schematically that when the actin cytoskeleton of a cell is disrupted by specific drugs, the cell envelope retracts, cylinder-like protrusions develop and a Rayleigh-Plateau instability can result.

Chapter 5

Membranes

The surfaces under tension discussed in the last chapter are the simplest examples for an effective interface model. We now turn to a case of large biological relevance, namely to lipid bilayers, which belong to the class of fluid membranes. For example, vesicles are simply a piece of lipid membrane closed into a spherical bag. The physics of fluid membranes is determined by their bending energy. Due to thermal fluctuations, membranes tend to make large excursions, which lead to steric interactions and scale-dependent renormalization of the bending rigidity. Finally we will explain how the concept of bending energy can be applied to predict vesicle shapes and adhesion.

5.1 Bending energy

In general, a thin elastic sheet made from isotropic material has three fundamental modes of deformation as shown in Fig. 5.1: in-plane shear, in-plane compression and out-of-plane bending. Examples for thin elastic sheet are sheet of paper, thin polymer films, single graphite layers and the cell envelope. *Fluid membranes* are thin films for which the shear modulus vanishes because the material flows under mechanical stress. The most prominent example for a fluid membrane is the lipid bilayer, which is the main structural element of biological cells. Since lipid bilayers consist of densly packed hydrocarbon chains, their compression modulus is rather large. Together with the vanishing shear modulus, this makes bending the only relevant deformation mode for fluid membranes. Because there are no natural coordinates for a fluid surface, the bending energy should be a scalar under coordinate

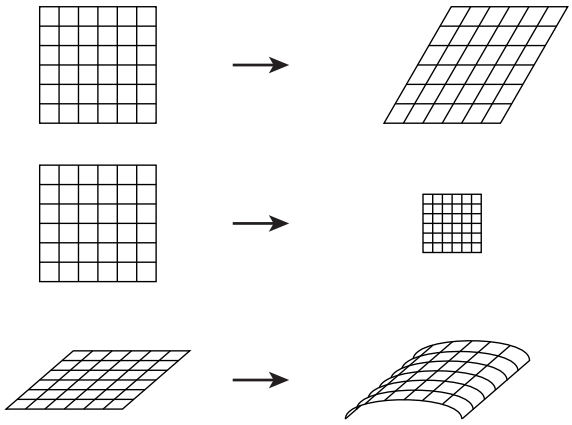


Figure 5.1: A thin film has three fundamental modes of deformation: in-plane shear, in-plane compression and out-off-plane bending.

transformations. Therefore it can depend only on the local curvatures. In a small curvature expansion, one thus gets the following *bending Hamiltonian* (also known as *Helfrich-Hamiltonian*):

$$H = \int dA \left\{ \sigma + 2\kappa(H - c_0)^2 + \bar{\kappa} K \right\} \tag{5.1}$$

The tension σ in lipid bilayers always has to be below the rupture tension of 10^{-3} J/m^2. For vesicles, osmotic deflation can be used to remove tension which has developed during vesicle preparation. In lamellar phases, usually no tension exists because the area per lipid can adjust to its optimal value. The *spontaneous curvature* $c_0 = 1/R_0$ results if a prefered radius of curvature exists. For lipid bilayers, such an asymmetry might result from adsorption of additional material to one side. For lipid monolayers, a spontaneous curvature results from hydration effects and therefore is a strong function of temperature. The *bending rigidity* κ describes the energy cost of cylindrical deformations. Typical values are k_BT for amphiphilic monolayers and $20k_BT$ for phospholipid bilayers. Due to the Gauss-Bonnet theorem $\int dAK = 2\pi\chi$ with the Euler characteristic χ describing the topology, the *Gaussian bending rigidity* $\bar{\kappa}$ is related to the energy cost of topology changes.

The bending rigidity κ has to be positive, otherwise an instability of spontaneous bending would occur. It can be measured experimentally in many different ways. In contrast, it is very hard to experimentally determine

the Gaussian bending rigidity $\bar{\kappa}$. In order to gain more insight into the physical meaning of $\bar{\kappa}$, we rewrite the Helfrich-Hamiltonian for the case $\sigma = 0$ and $c_0 = 0$ in the following way:

$$H = \int dA \left\{ \frac{1}{2}(\kappa + \frac{\bar{\kappa}}{2}) \, (\kappa_1 + \kappa_2)^2 - \frac{\bar{\kappa}}{4} \, (\kappa_1 - \kappa_2)^2 \right\} \tag{5.2}$$

From this formula, it follows that for a stable topology, one needs to have $-2\kappa < \bar{\kappa} < 0$. For $\bar{\kappa} < -2\kappa$, an instability with $\kappa_1 = \kappa_2 \to \infty$ towards infinitely many small droplets would result. For $\bar{\kappa} > 0$, an instability with $\kappa_1 = -\kappa_2 \to \infty$ towards a minimal surface with an infinitely small lattice constant would result. Therefore we conclude that $\bar{\kappa}$ should be small and negative.

$\bar{\kappa}$ is only relevant if topological changes can occur. Otherwise the Gauss-Bonnet theorem states that this term in the Helfrich Hamiltonian is a topological constant. Therefore for closed or periodic surfaces, we basically deal with solutions to the effective interface Hamiltonian $\int dAH^2$. These solutions are called *Willmore surfaces*. Since $H = 0$ is obviously a minimum of this Hamiltonian, minimal surfaces are a special subset of the set of solutions. In lipid-water mixtures, the most prominent phase therefore is the lamellar phase, for which both mean and Gaussian curvatures vanish. Another prominent example are the cubic phases, for which the midplane of the lipid bilayer has the shape of a triply periodic minimal surface with cubic symmetry (eg the gyroid). As already mentioned, minimal surfaces cannot be compact. However, Willmore surfaces can. Compact Willmore surfaces are eg the sphere ($G = 0$), the Clifford torus ($G = 1$) and the Lawson surface ($G = 2$). For a sphere, the Helfrich-Hamiltonian gives $H = 8\pi\kappa$, that is the bending energy is independent of size. This result in fact reflects the conformal invariance of the Helfrich-Hamiltonian, which for vesicles leads to the interesting property of conformal diffusion in shape space.

5.2 Thermal fluctuations

Like for the interfaces under tension, we first consider thermal fluctuations around a flat reference state. The bending Hamiltonian in the Monge para-

metrization reads

$$F = 2\kappa \int dA\ H^2 \approx 2\kappa \int dxdy\ \sqrt{1+h_x^2+h_y^2}\left(\frac{h_{xx}+h_{yy}}{2}\right)^2 \tag{5.3}$$

$$\approx \frac{\kappa}{2}\int dxdy\ (h_{xx}+h_{yy})^2 \tag{5.4}$$

In Fourier space this gives

$$\langle |h(\mathbf{q})|^2\rangle = \frac{k_BT}{\kappa q^4} \tag{5.5}$$

Going back to real space gives us

$$\langle h(\mathbf{r})^2\rangle = \frac{k_BT}{\kappa}\frac{1}{(2\pi)^2}\int_{2\pi/L} d\mathbf{q}\frac{1}{q^4} = \frac{1}{16\pi^3}\frac{k_BT}{\kappa}L^2 \tag{5.6}$$

where again the result is independent of position $\mathbf{r}$ due to translational invariance. In contrast to the result for an interface under tension, now a molecular cutoff is not needed. For system size $L = 10$ nm, we get a rms amplitude of 0.1 nm. For $L = 1$ cm, it is already 100 μm. Thus we conclude that a membrane makes much larger excursions than an interface under tension.

In order to study this issue in more detail, we can ask on which length scale membranes loose their orientation due to thermal fluctuations. In the Monge parametrization, the normal vector reads $\mathbf{n} = (-h_x, -h_y, 1)/\sqrt{\det g}$ with $\det g = 1+h_x^2+h_y^2$. Then we find for the angle θ between surface normal and z-axis:

$$\cos\theta \approx 1 - \frac{\theta^2}{2} = \mathbf{n}\cdot\mathbf{z} = \frac{1}{\sqrt{\det g}} \approx 1 - \frac{h_x^2+h_y^2}{2} \tag{5.7}$$

For the thermal average of the orientation we therefore have

$$\langle\theta^2\rangle = \langle(h_x^2+h_y^2)\rangle = \frac{k_BT}{(2\pi)^2\kappa}2\pi\int_{2\pi/L}^{2\pi/a} qdq\frac{q^2}{q^4} = \frac{k_BT}{2\pi\kappa}\ln(\frac{L}{a}) \tag{5.8}$$

Orientation is lost if $\langle\theta^2\rangle = \pi^2$, which defines a *membrane persistence length*

$$L_p = ae^{(2\pi^3\kappa)/(k_BT)} \tag{5.9}$$

also known as the *de Gennes-Taupin length*. For lipid bilayers, we have $\kappa = 20k_BT$ and L_p assumes the astronomical magnitude of ae^{400}. Therefore we conclude that although lipid bilayers make large excursions, they

do not loose orientation on realistic length scales and therefore the Monge parametrization is a reasonable approximation. For amphiphilic monolayers with $\kappa = 1 k_B T$, however, the persistence length L_p might become smaller than sample size. Then the membrane undergoes a *crumpling transition*, depending on the effect of self-avoidance, which cannot be treated in the Monge parametrization.

Up to now, we have worked in the Monge parametrization only in the lowest order of an expansion in small curvatures. If one includes higher orders, one has to go beyond simple Gaussian integrals and to use the technique of Feynmann diagrams. Then one finds that bending rigidity κ becomes length scale dependent:

$$\kappa(l) = \kappa - 3 \, \frac{k_B T}{4\pi} \, \ln\left(\frac{l}{a}\right) \tag{5.10}$$

The prefactor -3 has been confirmed in Monte Carlo simulations. The persistence length L_p can also be defined as the length l at which bending rigidity κ vanishes, giving essentially the same result as before.

5.3 Helfrich-interaction

Since membranes make large excursions, steric hinderance (eg by close-by walls or other membranes) leads to a large decrease in configurational entropy. This results in a long-ranged steric interaction, the so-called *Helfrich-interaction*. We first study this issue with scaling arguments. From Eq. (5.6) we know that the membrane makes one hump per distance L_p, with height h given by $h^2 \sim (k_B T/\kappa) L_p^2$. If this hump is suppressed by steric hinderance, the membrane looses entropy per area $\sim k_B/L_p^2$. The bending energy per area stored in the hump is $\sim \kappa (h/L_p^2)^2$, because h/L_p^2 is an approximation for the Laplace-term in Eq. (5.4). The overall change in free energy per area follows as

$$\Delta F \sim \kappa \left(\frac{h}{L_p^2}\right)^2 - T\left(\frac{-k_B}{L_p^2}\right) \sim \frac{(k_B T)^2}{\kappa} \frac{1}{h^2} \tag{5.11}$$

Therefore a long-ranged interaction $\sim 1/h^2$ with a strong temperature dependance results. Although a rigorous derivation of the Helfrich-interaction does not exist, it can be made more formal than in the simple scaling picture. Consider a membrane fluctuating in a confined space, eg between two walls.

The simplest way to model confinement is to introduce a harmonic potential:

$$\Delta F = \frac{\kappa}{2} \int dxdy \ \left\{ (h_{xx} + h_{yy})^2 + \frac{h^2}{\xi^4} \right\} \tag{5.12}$$

where the confinement length ξ is a parameter which should be adjusted in such a way that $\langle h^2 \rangle = \mu d^2$, with μ a geometrical constant of order unity and d the distance to the wall. The fluctuation amplitude can be calculated as above:

$$\langle h^2 \rangle = \frac{k_B T}{\kappa} \frac{1}{(2\pi)^2} 2\pi \int q \ dq \frac{1}{(q^4 + \xi^{-4})} = \frac{k_B T}{8\kappa} \xi^2 \tag{5.13}$$

Therefore

$$\xi = \left(\frac{8\mu\kappa}{k_B T} \right)^{1/2} d \tag{5.14}$$

The free energy difference is

$$\Delta F = \frac{k_B T}{(2\pi)^2} 2\pi \int q \ dq \ln \left(\frac{q^4 + \xi^{-4}}{q^4} \right) = \frac{k_B T}{8\xi^2} = \frac{(k_B T)^2}{64\kappa\mu d^2} \tag{5.15}$$

In terms of scaling, we thus get the same result as before, $\Delta F \sim (k_B T)^2/(\kappa d^2)$. However, we now also get the prefactors. Simple geometrical arguments and Monte Carlo simulations with triangulated surfaces give $\mu = 1/6$ and $\mu = 0.13$, respectively.

5.4 Vesicles

Vesicles are closed bags of lipid membrane. Experimentally, they come in different sizes (small, 20-50 nm, large, 50-500 nm, giant, 500 nm - 100 μm) and can be prepared with different techniques (ultrasonification, extrusion from a porous membrane, electroformation in an AC-field). The main technique to study vesicles is video microscopy, which reveals an amazing variety of different shapes. It also shows that their shape is constantly fluctuating due to thermal activation. Theoretically, vesicle shapes represented by mathematical surfaces can be investigated with different techniques. For the equilibrium shapes, one can either solve the Euler-Lagrange equations resulting from the bending Hamiltonian or use energy minimization techniques, either on some test parametrization or for triangulated surfaces. For fluctuations, one can

either look at expansions around some well-defined reference state or use Monte Carlo simulations of triangulated surfaces.

The physics of vesicles is determined by the bending energy introduced above in combination with three very important additional constraints. First, the area A of a vesicle can be considered to be constant, because the concentration of free lipids in the surrounding water is negligible (1 - 100 molecules per μm^3). This introduces a length scale R_0, namely the radius of the equivalent sphere, by $A = 4\pi R_0^2$. Second, also the volume V of a vesicle can be considered to be constant. Although the lipid bilayer is permeable for water molecules, this is not true for eg sugar molecules or ions, which are invariably present in vesicle experiments. Therefore any flux of water would set up an osmotic pressure which cannot be sustained by the vesicle and therefore is strongly suppressed. The combination of constant area A and constant volume V results in one important control parameter, the reduced volume $v = V/((4\pi/3)R_0^3)$. Because the sphere has the largest reduced volume possible, $v \leq 1$ always.

Third, the number of lipids in each monolayer can be taken to be constant, too, because the propability for change from one monolayer to the other is very low (for phospholipids, the timescale for flip-flop is hours, that is larger than typical labaratory time scales). If the number of lipids in the different leaflets is different due to the way the vesicle has formed, then this difference will stay. Since each lipid has a prefered molecular area, this leads to a prefered area difference ΔA_0 between inner and outer leaflet and therefore to a geometrical constraint. If the two leaflets are modelled as parallel surface with distance d, then the area difference can be calculated to be

$$\Delta A = 2d \int dAH \tag{5.16}$$

Therefore the larger ΔA, the stronger the vesicle will be curved. However, since this is a global constraint, it is not clear a priori where this curvature will be positioned. For the following, we non-dimensionalize the area difference to $m = \Delta A/4dR_0$. Derivation from the prefered value m_0 leads to a penalty term in the Hamiltonian, which can be shown to be harmonic. This leads to the *area-difference-elasticity (ADE) model*:

$$\frac{H}{\kappa} = 2 \int dAH^2 + \frac{\alpha}{2} (m - m_0)^2 \tag{5.17}$$

The stiffness of the area difference constraint α is a dimensionless material parameter of the order of unity. In fact it follows from a detailed derivation as

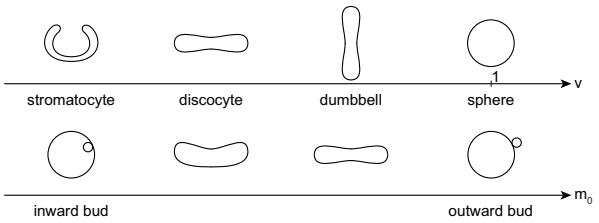


Figure 5.2: (a) Vesicle shapes as function of reduced volume v. Discocyte and stomatocyte shapes are known from red blood cells, but the echinocyte shape is missing. (b) Vesicle shapes as function of prefered area difference m_0. The echinocyte can be interpreted as a shape with many outward buds which are prevented from forming thin necks due to additional effects.

$\alpha = 2k_m d^2/\pi\kappa$, where k_m is the area compression modulus for a monolayer. The ADE-model defined by Eq. (5.17) together with the two constraints of constant area and constant volume has three parameters, namely reduced volume v, prefered area difference m_0 and the material parameter α.

One motivation for the extensive research of vesicle shapes is the hope to explain the shape of red blood cells. In 1974, it has been suggested by Sheetz and Singer that the stomatocyte - discocyte - echinocyte sequence which can be observed for changes in different variables (like pH, salt concentration, cholesterol concentration or ATP supply) can be explained by the effect of area difference (*bilayer couple hypothesis*). In order to compare this expectation with the physics of vesicles, we start by considering vesicle shapes without the area difference term. Then the only relevant parameter is reduced volume v, which for a sphere attains its maximally possible value 1. As v is reduced to smaller values, the sphere changes first into a dumbell, then into a discocyte and finally into a stomatocyte, compare Fig. 5.2. Although this sequence does contain two of the three shapes in the main sequence of red blood cells, it does not contain the echinocyte. If we now turn to the ADE-model, we find that outward and inward budded shapes are favorable for positive and negative values for the prefered area difference m_0, respectively. Although outward budded shapes resemble echinocytes, there are markingly different because the thin necks, which do not cost much bending energy ($H \approx 0$), are not present in echinocytes.

In order to obtain this shape, modelling red blood cells as vesicles is not sufficient. Instead, one has to include the elasticity of the polymer network underlying the plasma membrane (basically a spectrin network with mesh-

size 80 nm and an offset of 40 nm to the plasma membrane). At length scales above 100 nm, this system can be treated as a continuum, namely a thin (composite) shell surrounding a micron-sized bag of fluid (red blood cells do not have a nucleus and are filled with a concentrated solution of hemoglobin). The ADE-model is now augmented by a Hamiltonian for the in-plane elasticity of the polymer shell. Several different material laws have been suggested in the literature, for example

$$H = \int dA \left\{ \mu \frac{(\lambda_1 - \lambda_2)^2}{2\lambda_1\lambda_2} + \frac{K}{2}(\lambda_1\lambda_2 - 1)^2 \right\} \tag{5.18}$$

where μ and K are two-dimensional shear and bulk moduli, respectively, and $\lambda_i = 1 + u^{(i)}$ are the principal extension ratios. Note that in contrast to the vesicle case, a model for an elastic capsule now requires a reference state with an internal coordinate system, otherwise strain cannot be defined. For red blood cells, this reference shape seems to be oblate. For small deformations, Eq. (5.18) reduces to the linear isotropic elastic energy introduced before. For large deformations, Eq. (5.18) makes a particular choice, which is hard to test experimentally. However, the ADE-model Eq. (5.17) together with the elastic model Eq. (5.18) now is able to predict the echinocyte, where spicules now appear to be buds which are prevented from forming necks due to the elasticity of the polymer shell.

Chapter 6

Elasticity

On a short time scale, soft condensed matter under mechanical stress behaves like an elastic body, although on a long time scale, it might start to flow like a fluid. In this chapter we first introduce the fundamental concepts of elasticity theory, which are stress, strain and elastic moduli. In particular we specify them for the simplest elastic model, the *Hookean body*. Then we discuss some important applications of interest, including simple situations like pure shear as well as to more complicated situations like contact mechanics, which is important for e.g. AFM-measurements on cells.

6.1 Strain tensor

Elasticity theory is part of continuum mechanics, since it described deformations of condensed matter bodies on a scale which is much larger than the one of their building blocks. Here we are concerned with perfect elasticity, that is we assume that bodies resume their initial form completely after removal of the forces which cause the deformation.

Under deformation, the position vector of each body point is changed:

$$\mathbf{r} \to \mathbf{r}' = \mathbf{r} + \mathbf{u}(\mathbf{r}) \tag{6.1}$$

where $\mathbf{u}$ is the displacement vector field. We now consider two body points which are infinitely close to each other with distance dl:

$$dl^2 = dx_1^2 + dx_2^2 + dx_3^2 \tag{6.2}$$

In the following we make use of the Einstein convention (summation over repeated indices) and the Kronecker symbol:

$$dl^2 = dx_i^2 = dx_i dx_i = \delta_{ik} dx_i dx_k \tag{6.3}$$

The length element changes due to the deformation:

$$dl'^2 = (dx_i + du_i)^2 = (dx_i + \frac{\partial u_i}{\partial x_k} dx_k)^2 \tag{6.4}$$

$$= dl^2 + 2\frac{\partial u_i}{\partial x_k} dx_i dx_k + \frac{\partial u_i}{\partial x_k}\frac{\partial u_i}{\partial x_l} dx_k dx_l \tag{6.5}$$

$$= dl^2 + \frac{\partial u_i}{\partial x_k} dx_i dx_k + \frac{\partial u_k}{\partial x_i} dx_i dx_k + \frac{\partial u_l}{\partial x_k}\frac{\partial u_l}{\partial x_i} dx_i dx_k \tag{6.6}$$

$$= dl^2 + 2u_{ik} dx_i dx_k \tag{6.7}$$

where we have introduced the *strain tensor* u_{ik}:

$$u_{ik} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} + \frac{\partial u_l}{\partial x_k}\frac{\partial u_l}{\partial x_i}\right) \approx \frac{1}{2}\left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i}\right) \tag{6.8}$$

The last approximation is justified as long as $\partial u_i / \partial x_k \ll 1$. Since the strain tensor is symmetric, $u_{ik} = u_{ki}$, it can be diagonalized at every body point, with eigenvalues $u^{(i)}$. If we use the corresponding coordinates, then

$$dl'^2 = (1 + 2u^{(i)}) dx_i^2 \tag{6.9}$$

Thus the overall deformation can be decomposed into three independent ones, which correspond to stretching/compressing the body along the three orthogonal coordinate directions. The relative changes in length along these directions are

$$\frac{\sqrt{1 + 2u^{(i)}} dx_i - dx_i}{dx_i} = \sqrt{1 + 2u^{(i)}} - 1 \approx u^{(i)} \tag{6.10}$$

Therefore the eigenvalues of the strain tensor represent the relative changes in length caused by the deformation.

The change in volume follows as

$$dV' = dV(1 + u^{(1)})(1 + u^{(2)})(1 + u^{(3)}) \tag{6.11}$$

$$\approx dV(1 + u^{(1)} + u^{(2)} + u^{(3)}) = dV(1 + u_{ii}) \tag{6.12}$$

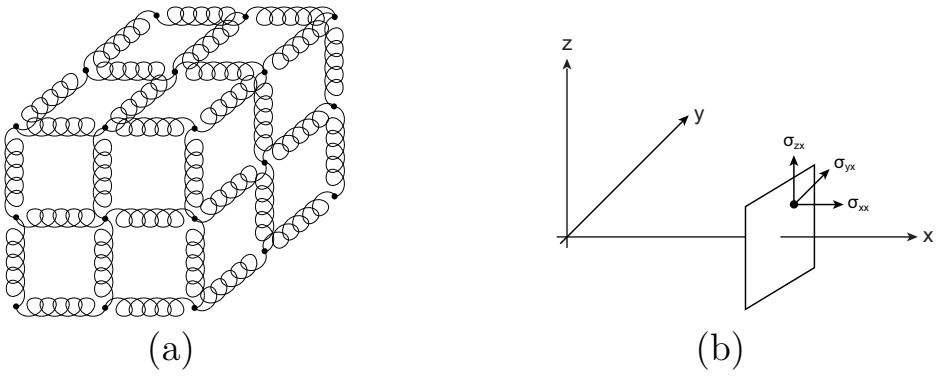


Figure 6.1: (a) Linear elasticity assumes that force is transmitted locally, like through an array of springs. (b) Definition of the stress tensor: σ_{ij} is the i-th component of the force acting on the area element A_k.

Note that the trace of the strain tensor, u_{ii}, is invariant under coordinate transformations. Physically, it corresponds to the relative change in volume caused by the deformation:

$$\frac{dV' - dV}{dV} = u_{ii} \tag{6.13}$$

Deformations with no volume change are called *pure shear*, $u_{ii} = 0$. The opposite case is *homogeneous compression*, in which the body shape does not change, $u_{ik} \sim \delta_{ik}$. Each deformation can be decomposed into a shear and a compression part:

$$u_{ik} = \left(u_{ik} - \frac{1}{3}\delta_{ik} u_{ll} \right) + \frac{1}{3}\delta_{ik} u_{ll} \tag{6.14}$$

6.2 Stress tensor

The strain tensor describes the effect of a deformation. We now turn to its cause, the forces. Under deformation, internal forces develop in the body, which in elasticity theory are assumed to be short-ranged (imagine beads connected by springs). Therefore they can act on a certain volume of a body only over its surface. If the sum over all forces on a certain volume element can be written as an integral over its surface, then there must exist a tensor

of rank two, the so-called *stress tensor* σ_{ik}, such that

$$\int F_i dV = \int \sigma_{ik} dA_k = \int \frac{\partial \sigma_{ik}}{\partial x_k} dV \tag{6.15}$$

$\sigma_{ik} dA_k$ (no summation) is the i-th component of the force acting on the area element dA_k *from outside*. For example, there is a force σ_{xx} acting normal to the unit area which is orthogonal to the x-axis; the forces acting tangentially on the same area are σ_{yx} and σ_{zx}. Note that $-\sigma_{ik} dA_k$ is the force acting *from inside*. In equilibrium, it has to cancel the force acting from outside. There is a certain gauge freedom for the stress tensor which allows to write it symmetric in any case, $\sigma_{ik} = \sigma_{ki}$. This ensures that also the angular moment on a volume can be written as a surface integral.

In equilibrium without deformation, there are no internal forces and $\sigma_{ik} = 0$. In equilibrium with deformation, all internal forces have to balance:

$$\frac{\partial \sigma_{ik}}{\partial x_k} + f_i = 0 \tag{6.16}$$

where f_i are volume forces like gravity, ρg_i. Usually deformations are caused by forces P_i acting on the body surface. They enter here as boundary conditions:

$$P_i dA - \sigma_{ik} dA_k = P_i dA - \sigma_{ik} n_k dA = 0 \tag{6.17}$$

where n_k is the surface normal, thus

$$P_i - \sigma_{ik} n_k = 0 \tag{6.18}$$

has to be satisfied on the body surface. Special boundary conditions are free ($P_i = 0$, that is $\sigma_{ik} n_k = 0$, no surface stresses) or clamped ($u_i = 0$, displacement fixed at boundary).

6.3 Elastic moduli

The free energy of a deformed body is a scalar and in linear elasticity theory has to be a quadratic function of the strain tensor, thus

$$F = \frac{1}{2} C_{iklm} u_{ik} u_{lm} \tag{6.19}$$

Since the strain and stress tensors are symmetric, the tensor of the elastic moduli has to have the following symmetries

$$C_{iklm} = C_{kilm} = C_{ikml} = C_{lmik} \tag{6.20}$$

It can be shown that such a tensor has 21 independent components. The actual number for a given material depends on the crystal system. For example, it is 21 for the triclinic system, 5 for the hexagonal system (like graphite) and 3 for the cubic system.

For an isotropic elastic body, there are only two elastic moduli, because there are only two ways to contract the strain tensor in second order into a scalar:

$$F = \mu u_{ik}^2 + \frac{\lambda}{2} u_{ll}^2 = \mu \left(u_{ik} - \frac{1}{3}\delta_{ik} u_{ll} \right)^2 + \frac{K}{2} u_{ll}^2 \tag{6.21}$$

This defines two equivalent pairs of elastic constants, (μ, λ) and (μ, K). The first ones are called Lamé coefficients and correspond to $C_{iklm} = (\lambda/2)\delta_{ik}\delta_{lm} + \mu\delta_{il}\delta_{km}$. In the second term of Eq. (6.21), we have used the decomposition into shear and compression. Therefore μ is called the shear modulus and $K = \lambda + 2\mu/3$ the compression modulus; these two parameters have to be positive for thermodynamic stability. Since the isotropic is the simplest case, there is a minimum of two elastic constants for any material.

We now calculate the work per volume done by the internal forces against a small change in displacement:

$$\frac{\partial \sigma_{ik}}{\partial x_k} \delta u_i = -\sigma_{ik} \frac{\partial \delta u_i}{\partial x_k} = -\sigma_{ik} \frac{1}{2} \left(\frac{\partial \delta u_i}{\partial x_k} + \frac{\partial \delta u_k}{\partial x_i} \right) = -\sigma_{ik} \delta u_{ik} \tag{6.22}$$

where we have used a partial integration, neglected the boundary term at infinity (where the stress tensor is assumed to vanish), and used that the stress tensor is symmetric. We now can conclude that the elastic contribution to the free energy is

$$dF = \sigma_{ik} du_{ik} \tag{6.23}$$

and that for the isotropic case the stress tensor follows from the free energy as

$$\sigma_{ik} = \frac{\partial F}{\partial u_{ik}} = 2\mu u_{ik} + \lambda \delta_{ik} u_{ll} \tag{6.24}$$

This equation can easily be inverted. First we note from Eq. (6.24) that

$$\sigma_{ll} = (2\mu + 3\lambda) u_{ll} = 3K u_{ll} \tag{6.25}$$

The two traces are related to each other only by the compression modulus. Using again Eq. (6.24), we have

$$u_{ik} = \frac{1}{2\mu}\sigma_{ik} - \frac{\lambda}{2\mu(3\lambda+2\mu)}\delta_{ik}\sigma_{ll} \tag{6.26}$$

Thus the strain tensor is a linear function of the stress tensor. In other words, the deformation is proportional to the forces applied. This is simply the tensorial form of *Hooke's law* (Robert Hooke wrote in 1678: Ut tensio sic vis, that is Force is proportional to elongation).

The relation between stress tensor and strain tensor, Eq. (6.24), and the definition of strain, Eq. (6.8), allow to rewrite the equilibrium condition from Eq. (6.16) in terms of displacement:

$$\frac{\partial\sigma_{ik}}{\partial x_k} + f_i = \mu\frac{\partial^2 u_i}{\partial x_k^2} + (\mu+\lambda)\frac{\partial^2 u_k}{\partial x_i \partial x_k} + f_i = 0 \tag{6.27}$$

In vector notation, this equation reads

$$\mu\Delta\mathbf{u} + (\mu+\lambda)\nabla(\nabla\cdot\mathbf{u}) + \mathbf{f} = \mathbf{0} \tag{6.28}$$

Note that this equation is similar to, but more complicated than the Laplace equation. In particular, the second term couples the different components of the displacement vector $\mathbf{u}$. If boundary conditions can be written in terms of displacement, one has to solve this differential equation. If they are written in terms of stress, one has to make sure that the stress and strain tensors found are compatible with a displacement field.

Due to linearity, displacement follows from force by convolution with a Green function

$$u_i(\mathbf{r}) = \int d\mathbf{r}' G_{ik}(\mathbf{r}-\mathbf{r}')F_k(\mathbf{r}') \tag{6.29}$$

where the Green function G_{ik} follows from solving Eq. (6.28) for a point-like force at the origin:

$$\{\mu\Delta + (\mu+\lambda)\nabla\nabla\cdot\}\,\mathbf{u} = -\mathbf{F}\delta(\mathbf{r}) \tag{6.30}$$

For the infinite isotropic elastic space, the result is the one given by Kelvin in 1848:

$$G_{ik} = \frac{1}{8\pi\mu(2\mu+\lambda)}\left\{(3\mu+\lambda)\delta_{ik} + (\mu+\lambda)\frac{x_i x_k}{r^2}\right\}\frac{1}{r} \tag{6.31}$$

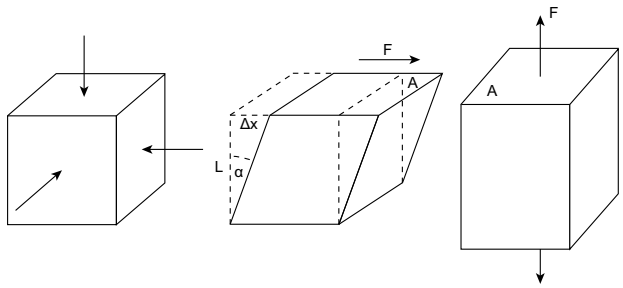


Figure 6.2: Three simple deformation experiments: (a) pure compression, (b) pure shear and (c) uniaxial stretch.

Although we assumed that internal forces are only short-ranged, elastic effects propagate with $1/r$ and thus are long-ranged. It is quite remarkable that such a strong result follows from relatively few and reasonable assumptions. In this sense, elasticity theory reminds of thermodynamics, which also is very general because it is based on a few reasonable assumptions.

In summary, linear isotropic elasticity theory involves 15 unknown (u_i, u_{ik}, σ_{ik}), which can be determined from the 15 equations resulting from Eq. (6.8) (definition of strain tensor), Eq. (6.16) (condition for mechanical equilibrium) and Eq. (6.24) (relation between stress and strain).

6.4 Some simple cases

6.4.1 Pure (hydrostatic) compression

Here a normal force towards the body interior acts on every volume element:

$$-pdA_i = -p\delta_{ik}dA_k = \sigma_{ik}dA_k \tag{6.32}$$

therefore

$$\sigma_{ik} = -p\delta_{ik} \tag{6.33}$$

The free energy then follows as

$$dF = \sigma_{ik}du_{ik} = -pdu_{ll} = -pdV \tag{6.34}$$

as it should. Since $\sigma_{ll} = 3Ku_{ll}$, we also find

$$u_{ll} = \frac{\sigma_{ll}}{3K} = -\frac{3p}{3K} = -\frac{p}{K} \tag{6.35}$$

that is

$$\frac{1}{K} = -\frac{1}{V}\frac{\partial V}{\partial p} \tag{6.36}$$

Therefore $1/K$ is the isothermal compressibility.

6.4.2 Pure shear

We shear a rectangular body with tangential stress $s = F/A$ on its upper side with the lower side being fixed. Then $\sigma_{12} = s$. It follows that $u_{12} = s/2\mu$ and $\mathbf{u} = (sy/\mu, 0, 0)$. The angle of deformation will be $\alpha = u_1/y = s/\mu$. This explains why μ is called *shear modulus*.

6.4.3 Uniaxial stretch

Consider a dilation in z-direction of a rectangular slab due to forces per area p acting on the upper and lower surfaces. Both strain and stress tensor will be constant. At the sides, no forces act, thus there $\sigma_{ik}n_k = 0$. Thus only σ_{zz} is non-vanishing. At the upper and lower surface, we have $\sigma_{zk}n_k = \sigma_{zz} = p$. The strain tensor follows from Eq. (6.26):

$$u_{zz} = \frac{(\lambda+\mu)p}{\mu(3\lambda+2\mu)} = \frac{p}{E} \tag{6.37}$$

$$u_{xx} = u_{yy} = -\frac{\lambda p}{2\mu(3\lambda+2\mu)} = -\nu u_{zz} = -\frac{\nu p}{E} \tag{6.38}$$

Here we have defined two new elastic constants: *Young modulus* E and *Poisson ratio* ν:

$$E = \frac{\mu(3\lambda+2\mu)}{(\lambda+\mu)},\ \nu = \frac{\lambda}{2(\lambda+\mu)} \tag{6.39}$$

u_{zz} describes the relative change in length, and u_{xx} and u_{yy} describe lateral contraction. Since μ and K both have to be positive, $-1 < \nu < 1/2$. For most materials, ν is between 0 and 1/2. However, material with negative Poisson ratio is also known, in this case the sample expands in the lateral direction when being pulled (the simplest example for this case is crumpled paper). The relative change in volume is

$$u_{ii} = \frac{(1-2\nu)p}{E} \tag{6.40}$$

thus $\nu = 1/2$ corresponds to the incompressible case (eg rubber). For the displacement field we find

$$\mathbf{u} = \frac{p}{E}(-\nu x, -\nu y, z) \tag{6.41}$$

and the free energy is

$$F = \frac{p^2}{2E} \tag{6.42}$$

This form of the energy reminds of the harmonic spring. Indeed the elastic slab under uniaxial stretch has an effective spring constant $k = EA/L$.

Note that in a real stretch experiment, one usually has displacement rather than stress boundary conditions (usually the slap is clamped at both sides and then the clamps are moved apart), but there is a general theorem in elasticity theory (*St. Venant's principle*) that states that the details of the boundary do not matter as long as they are farer apart then the largest dimension over which the load is applied.

If stress is applied simultaneously to different faces, the solution simply follows by superposition, e.g.

$$u_{zz} = \frac{\sigma_{zz}}{E} - \frac{\nu\sigma_{xx}}{E} - \frac{\nu\sigma_{yy}}{E} \tag{6.43}$$

6.5 Beams and thin shells

For thin shells made from isotropic linear material, there are three fundamental modes of deformation: in-plane shear, in-plane compression and out-of-plane bending.

6.6 Capsules

We consider a thin plate of thickness h which is bent into a sphere of radius R in its unstressed state (shell). Consider a radial expansion by ζ. Then the change in curvature is ζ/R^2 and the change in strain is ζ/R. The elastic constants for bending and stretching are Eh^3 and Eh, respectively, there E is Young modulus. Therefore we have

$$F_b = Eh^3(\frac{\zeta}{R^2})^2, F_s = Eh(\frac{\zeta}{R})^2, \frac{F_s}{F_b} = (\frac{R}{h})^2 \gg 1 \tag{6.44}$$

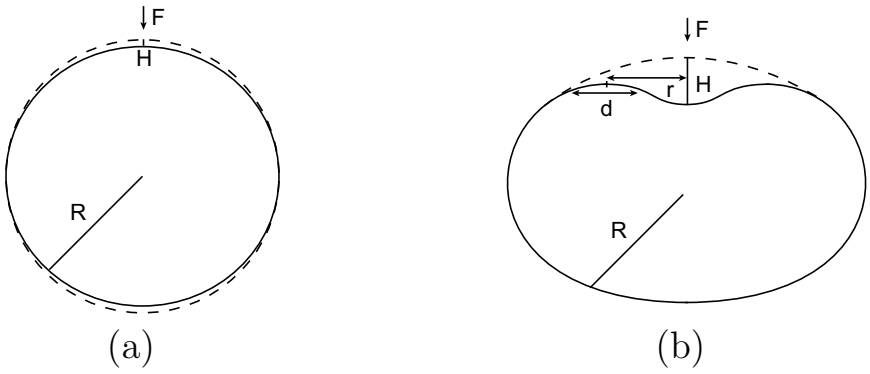


Figure 6.3: Capsule deformation. (a) For small force, the capsule behaves like a linear spring. (b) For large force, inward buckling occurs.

and stretching is much more expansive then bending. Usually these systems avoid stretching, eg when a cylinder with open ends is deformed. However, if a sphere is deformed, stretching cannot be avoided. In membrane theory, one neglects the bending energy and considers only stretching.

Now consider a sphere of radius R which is indented by a force f over a lateral distance d. The elastic energy $F = F_b + F_s$ has contributions from bending F_b and stretching F_s:

$$F_b = Eh^3(\frac{\zeta}{d^2})^2 d^2, F_s = Eh(\frac{\zeta}{R})^2 d^2 \tag{6.45}$$

where ζ/d^2 is change in curvature and ζ/R is strain. F_b decreases with d, while F_s increases with d. The balance occurs at

$$d \sim \sqrt{Rh}, F \sim \frac{Eh^2}{R}\zeta^2 \tag{6.46}$$

Since $F = f\zeta$, the indentation ζ resulting from force f is

$$\zeta \sim \frac{fR}{Eh^2} \tag{6.47}$$

Thus the spring constant is Eh^2/R.

If the force becomes larger, the shell can buckle inwards and the elastic energy becomes localized along a ring of lateral size d and radius r. The geometrical relation between indentation H and ring radius r is $r = \sqrt{RH}$.

We now express the elastic energies in terms of r:

$$F_b = Eh^3(\frac{\zeta}{d^2})^2(rd), F_s = Eh(\frac{\zeta}{R})^2(rd) \tag{6.48}$$

where now $\zeta = rd/R$. Minimization for d again yields $d \sim \sqrt{Rh}$. Thus the overall energy now is

$$F \sim Er^3(\frac{h}{R})^{5/2} = \frac{Eh^{5/2}}{R}H^{3/2} \tag{6.49}$$

Note that there is a size-dependent line tension here, $F/r \sim r^2$. Comparing Eq. (6.46) and Eq. (6.49) shows that buckling becomes more favorable at $H \sim h$, that is for

$$f \sim \frac{Eh^3}{R} \tag{6.50}$$

Since for large indentation H buckling occurs and the elastic energy scales only with $H^{3/2}$ rather than with H^2, there exists a buckling instability with regard to pressure p, since pressure has a stronger scaling with H:

$$F_p \sim -pr^2H \sim -pRH^2 \tag{6.51}$$

The critical indentation, beyond which buckling grows, follows by adding Eq. (6.49) and Eq. (6.51) and finding the maximum:

$$H_c \sim \frac{Eh^5}{R^4p^2} \tag{6.52}$$

Since buckling occurs for $H \sim h$, we can estimate the critical pressure by setting $H_c = h$ and find

$$p_c \sim E(\frac{h}{R})^2 \tag{6.53}$$

6.7 Contact mechanics

Contact mechanics describes the elastic deformations which occur when two bodies are pressed onto each other. Examples include two colliding elastic spheres, a sphere pressed onto a substrate by gravity, and an AFM-tip indenting a sample. In the latter case, one usually measures the force-indentation curve. Force-indentation curves depend sensitively on geometry. The case

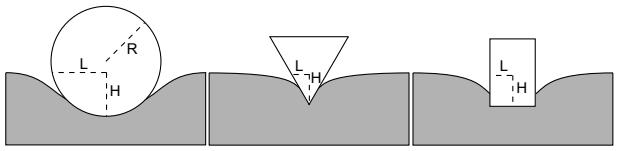


Figure 6.4: Contact mechanics: a hard object (e.g. the tip of an AFM) indents a soft substrate by distance H. (a) Spherical indenter. (b) Conical indenter. (c) Flat indenter.

of a spherical indenter has been solved by Hertz in 1881. The case of an arbitrary axisymmetric indenter has been solved independently by Sneddon in 1965 and Ting in 1966. In 1971, Johnson, Kendall and Roberts (JKR) extended the Hertz analysis to include adhesion.

Consider an indenter which is pressed by a force F onto an elastic half-space and indents it by h. The force-indentation curve will depend on the indenter's geometry. We consider three different geometries, the spherical, conical and flat cylindrical indenters, which all have radial symmetry. There are two general statements which we can make before doing any calculation. First, the vertical displacement will fall off like $1/r$ in the far field. Second, from the origin to some contact radius a, the elastic material has to follow the shape given by the rigid indenter. However, the general problem is hard to solve, in particular because we deal with mixed boundary conditions (fixed displacement in the contact region, free surface away from it).

In order to obtain the force-indentation-relations, we first consider a scaling approach. The elastic energy scales as $U \sim VEu^2$, where V is the volume of the part of the elastic medium which is deformed considerably, E is an elastic modulus and u is dimensionless strain. For the spherical indenter with radius R, indentation by h results in a deformation of lateral size $l \sim \sqrt{Rh}$. With $V \sim l^3$ and $u \sim h/l$, we find $U \sim ER^{1/2}h^{5/2}$ and thus $F = dU/dh \sim ER^{1/2}h^{5/2}$. For a conical indenter, $l = h/\tan\alpha$ where α is the angle between the substrate and the cone. Again $V \sim l^3$ and $u \sim h/l$, thus $U \sim Eh^3/\tan\alpha$ and $F \sim Eh^2/\tan\alpha$. This result is consistent with the fact that a cone does not have any intrinsic length scale, hence $\sqrt{F/E}$ is the only one present. For the flat cylindrical indenter, appreciable strain can only occur near the indenters sharp edge. The edge length scales as the cylinder radius a. The length over which the deformation decays from the

edge has to be independent of a, thus it has to scale with h. Hence we have $V \sim ah^2$ and $u \sim 1$. Therefore $U \sim Eah^2$ and $F \sim Eah$. Note that the flat indenter is the only one for which force F scales linearly with indentation h. At the same time we work within linear elasticity theory and use Hooke's law. Therefore the non-linearity of the force-indentation curves are purely of geometrical origin.

Comparision with the full solution shows that the scaling approach was correct ($M = E/(1-\nu^2)$):

- Sphere with radius R:

$$H = \left(\frac{3}{4}\right)^{2/3} \frac{1}{R^{1/3}} \left(\frac{F}{M}\right)^{2/3} \tag{6.54}$$

- Cone with slope α:

$$H = \left(\frac{\pi}{2}\right)^{1/2} \tan\alpha \left(\frac{F}{M}\right)^{1/2} \tag{6.55}$$

- Cylindrical slab with radius a:

$$H = \frac{1}{2a} \left(\frac{F}{M}\right) \tag{6.56}$$

Finally we discuss the JKR-theory for elastic contact of a sphere in the presence of adhesion. For the contact radius a, the Hertz-model gives

$$a^3 = \frac{3}{4M} RF \tag{6.57}$$

With a finite adhesion energy W per unit area, this relation changes to

$$a^3 = \frac{3}{4M} R \left\{ F + 3\pi RW + \sqrt{6\pi RWF + (3\pi RW)^2} \right\} \tag{6.58}$$

Therefore at vanishing loading ($F = 0$), a finite contact radius with $a_0^3 = (9\pi R^2 W)/(2M)$ emerges. In order to detach the body, a critical force $F_c = -3W\pi R/2$ is required, which surprisingly is independent of the elastic constants. The critical contact radius turns out to be $a_c = a_0/4^{1/3} = 0.63 a_0$.

Chapter 7

Hydrodynamics

Hydrodynamics describes the flow of simple fluids, in contrast to the flow of complex fluids, which are described by rheology. In this chapter, we introduce the concept of a *Newtonian fluid* and apply it to some simple cases of interest. Flow of Newtonian fluids is described by the non-linear Navier-Stokes equation, but for most biological applications, one only needs the special case of the linear Stokes equation (small Reynolds number regime), because at the small length scale of biomolecules and cells, inertia is irrelevant. Because life at small Reynolds number is very different from our own life at large Reynolds number, intuition might fail when discussing transport issues at the scale of cells.

7.1 Fundamentals

The fundamental equations of fluid mechanics follow from conservation laws. For most situations, one has to consider conservation of mass and momentum. If there are no sinks or sources for mass, changes inside a given volume can only occur by flux $j = \rho v$ over its boundaries, where ρ is density and $\mathbf{v}$ is velocity:

$$\frac{\partial}{\partial t}\int_V \rho dV = \int_V \frac{\partial \rho}{\partial t} dV = -\int_{\partial V} (\rho \mathbf{v}) dA = -\int_V \nabla(\rho \mathbf{v}) dV \qquad (7.1)$$

where we have used Gauss' theorem. Since this relation has to hold for any volume V, we arrive at the *equation of continuity*:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \qquad (7.2)$$

For incompressible fluids, density ρ is constant and therefore $\nabla \cdot \mathbf{v} = 0$.

In regard to conservation of momentum mv, we can make use of Newton's 2nd law:

$$\rho \frac{d\mathbf{v}}{dt} = \rho \left[\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} \right] = \nabla \cdot \sigma + \mathbf{f} \tag{7.3}$$

where we have assumed that ρ is constant. On the left hand side we have the total derivative of velocity for time (also called *substantial derivative*), which follows from $d\mathbf{v} = \mathbf{v}(\mathbf{r} + \mathbf{v}dt, t + dt) - \mathbf{v}(\mathbf{r}, t)$. σ is the stress tensor and its gradient describes internal forces, e.g. due to shear. $\mathbf{f}$ describes external body forces like gravity.

In linear approximation, the stress tensor is proportional to the rate of the deformation tensor (that is to the time derivative of the strain tensor). For isotropic material, one has

$$\sigma_{ik} = -\delta_{ik} p + \eta \left(\frac{\partial v_i}{\partial x_k} + \frac{\partial v_k}{\partial x_i} - \frac{2}{3} \delta_{ik} (\nabla \cdot \mathbf{v}) \right) + \zeta \delta_{ik} (\nabla \cdot \mathbf{v}) \tag{7.4}$$

The first term represents the state of stress if the fluid is at rest, then pressure p is the only way to create some internal force. The two parameters η and ζ are the analogues of the shear and bulk moduli in elasticity theory and are called *shear* and *bulk viscosities*. Note that like in elasticity, the shear contribution is made traceless (no change in volume). Fluids described by this linear relation are called *Newtonian fluids*. Non-Newtonian fluids have viscosities which depend on shear rate and history (e.g. polymer melts and many biological materials like blood).

Using this form of the stress tensor in Eq. (7.3) yields:

$$\rho \left[\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} \right] = -\nabla p + (\zeta + \frac{1}{3}\eta) \nabla (\nabla \cdot \mathbf{v}) + \eta \Delta \mathbf{v} + \mathbf{f} \tag{7.5}$$

Note that the equation is non-linear due to the inertia terms. It has to be solved always together with the equation of continuity. For an incompressible fluid, $\nabla \cdot \mathbf{v} = 0$ and we obtain the *Navier-Stokes equation*:

$$\rho \left[\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} \right] = -\nabla p + \eta \Delta \mathbf{v} + \mathbf{f} \tag{7.6}$$

The Navier-Stokes equation and the equation of continuity, $\nabla \mathbf{v} = 0$, yield four equations for the four unknowns $\mathbf{v}, p$.

There are two important special cases:

- Inviscid flow: viscosity parts can be neglected, *Euler equation*:

$$\rho\left[\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v}\right] = -\nabla p + \mathbf{f} \tag{7.7}$$

  In the special case of potential (or irrotational) flow, $\nabla \times \mathbf{v} = \mathbf{0}$ and $\mathbf{v} = \nabla\Phi$, the equation of continuity becomes the Laplace equation $\Delta\Phi = 0$. For incompressible inviscid flow, rotations cannot develop, thus if flow is irrotational, it will remain irrotational.

- Creeping flow: inertia parts can be neglected, *Stokes equation*:

$$\eta\Delta\mathbf{v} = \nabla p - \mathbf{f} \tag{7.8}$$

  In this case, it is usually pressure which drives the flow.

Since the inertia terms scale as $\rho v^2/L$ and the viscosity terms as $\eta v/L^2$, where L is a typical length of the system, the relative magnitude of inertia and viscosity can be described by a dimensionless number, the *Reynolds number* Re:

$$Re = \frac{\rho v L}{\eta} \tag{7.9}$$

For $Re \ll 1$ and $Re \gg 1$, Stokes and Euler flow, respectively, applies. For water, $\rho = g/cm^3$. At $20°C$, water has $\eta = 10^{-3}Pas = 1cP$. Blood is not a Newtonian fluid, but blood plasma is, with a viscosity essentially that of water. Glycerin has $\eta = 1500cP$, a cell $\eta = 10^7 cP$, and glass $\eta > 10^{15}cP$. Turbulence usually occurs for large Re, but for special cases has been observed down to values in the order of 10; otherwise flow is laminar.

It is instructive to compare the Reynolds numbers for fish and cells. For a big and fast swimming fish, we have

$$Re = \frac{(g/cm^3)(m/s)m}{cP} = 10^6 \tag{7.10}$$

For a cell in blood flow in capillary venules we have

$$Re = \frac{(g/cm^3)(mm/s)(10\mu m)}{cP} = 10^{-2} \tag{7.11}$$

Therefore the fish and the cell live in very different hydrodynamic worlds. The fish world is Newtonian (acceleration vanishes with force, but inertia will

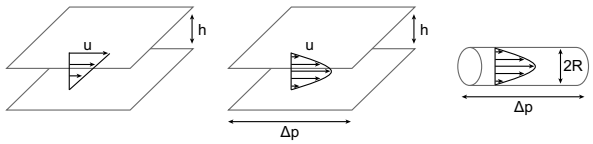


Figure 7.1: Three simple flow situations. (a) Linear shear flow between two parallel walls with relative motion (Couette flow). (b) Pressure-driven flow between two parallel walls (plane Poiseuille flow). (c) Pressure-driven flow in a capillary (cylindrical Poiseuille flow).

keep it going), the cell world is Aristotelian (velocity vanishes with force, the cell will stop if force is removed). Being a cell means living in a syrup.

In the presence of surfaces, usually one has to implement *no-slip boundary conditions*, that is velocity should vanish at the walls, due to attractive interactions between walls and fluid. This implies that Re is always small near surfaces. In order to deal with this problem in Euler flow, Prandtl has introduced the concept of *boundary layers*, that is thin layers in which the effect of viscosity is important no matter how high Re might be in the bulk (although its thickness vanishes as $Re \to \infty$). If outside the boundary layers one has potential flow, then rotation (turbulence) can only be generated by the boundary layer.

7.2 Simple examples

We now consider steady state solutions (no partial derivative for time) of the Navier-Stokes equation for situations in which the non-linear term disappears because there is no variation in velocity along the direction of flow. Therefore the same solutions also result if one starts from the Stokes equation.

7.2.1 Shear (Couette) flow

Two parallel surface with distance h are placed perpendicular to the z-direction. The lower one is held fixed, while the upper one is moved in x-direction with velocity u. By symmetry, $v_y = v_z = 0$ and $v_x = v_x(z)$. In this situation, the pressure gradient vanishes and the Stokes equation reads

$$\frac{\partial^2 v_x}{\partial z^2} = 0 \tag{7.12}$$

The solution with the correct boundary conditions is simply

$$v_x = \frac{zu}{h} \tag{7.13}$$

so the flow profile is linear. This result provides an operational definition for the viscosity: the external force has to balance the internal force, which from the above definition of the stress tensor is

$$F = \eta A \left(\frac{\partial v_x}{\partial z} \right)_{z=h} \tag{7.14}$$

thus

$$\eta = \frac{Fh}{Au} \tag{7.15}$$

Shear stress is defined as

$$\tau = \frac{F}{A} = \eta \frac{dv_x}{dz} = \eta \dot{\gamma} \tag{7.16}$$

where $\dot{\gamma} = dv_x/dz$ is called *shear rate*. The volume flowing through the channel per time is

$$Q = \frac{\partial V}{\partial t} = L_y \int_0^h v_x dz = L_y \frac{1}{2} uh \tag{7.17}$$

where L_y is the width of the channel.

7.2.2 Plane Poiseuille flow

Two parallel surface with distance h are placed perpendicular to the z-direction. Fluid flows along the positive x-direction driven by a homogeneous pressure gradient $\partial p/\partial x = -G$ (note that the gradient has to be negative for the fluid to flow to the positive z-direction). By symmetry, $v_y = v_z = 0$ and $v_x = v_x(z)$. The Stokes equation now read

$$\eta \frac{\partial^2 v_x}{\partial z^2} = -G \tag{7.18}$$

The solution which disappears at the boundary is

$$v_x(z) = \frac{G}{8\eta} \left[h^2 - 4(z - \frac{h}{2})^2 \right] = \frac{G}{2\eta} z(h - z) \tag{7.19}$$

and the flow profile is parabolic. The volume flow rate is

$$Q = L_y \int_0^h v_x dz = L_y \frac{Gh^3}{12\eta} \tag{7.20}$$

thus showing a strong scaling with h.

7.2.3 Cylindrical Poiseuille flow

We use cylindrical coordinates: the pipe direction is taken to be the z-direction, and perpendicular to it we have the radial variable r. By symmetry, the velocity field can only have a z-component $v_z = v_z(r)$. The Stokes equation now reads

$$\eta \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} \right) v_z = -G \tag{7.21}$$

where G is the pressure gradient in positive z-direction (note that the gradient has to be negative for the fluid to flow to the positive z-direction). The solution which disappears at the boundary is

$$v_z(r) = \frac{G}{4\eta}(R^2 - r^2) \tag{7.22}$$

The volume flowing per time through the crossection is

$$Q = 2\pi \int_0^R dr\ r\ v_z(r) = \frac{\pi G R^4}{8\eta} \tag{7.23}$$

This result is also known as the *Hagen-Poiseuille law*. Note the very strong scaling with R.

7.3 Stokes flow

We now turn to some important results which result only if one start from the Stokes equation for viscous flow.

7.3.1 Hydrodynamic lubrication

In lubrication, one uses thin liquid films to prevent contact of fast moving bodies. Consider a body with a mildly sloped surface, $h(x) = h_0 - \alpha x$,

which extends from $x = 0$ to $x = L$. The lower surface moves to the right with velocity V. Since h is much smaller than L, pressure variation across the gap can be neglected, $p = p(x)$. Thus the lubrication approximation is assuming that locally one has Couette-Poiseuille flow. Velocity only varies in z-direction, thus the Stokes equation reads

$$\eta \frac{\partial^2 v_x}{\partial z^2} = \frac{\partial p}{\partial x} \tag{7.24}$$

which can be integrated to

$$v_x = V\frac{(h-z)}{h} - \frac{1}{2\eta}\frac{\partial p}{\partial x} z(h-z) \tag{7.25}$$

which is obviously a superposition of Couette and Poiseuille flow. Correspondingly, the volume flow rate at position x reads

$$Q(x) = L_y(\frac{1}{2}Vh(x) - \frac{\partial p}{\partial x}\frac{h(x)^3}{12\eta}) \tag{7.26}$$

Conservation of mass requires Q not to depend on x and we can solve this equation for the pressure gradient:

$$\frac{\partial p}{\partial x} = \frac{6\eta V}{h(x)^2} - \frac{12\eta Q}{h(x)^3} \tag{7.27}$$

As boundary conditions, we set $p_0 = p_L = p_a$, the ambient pressure. We determine Q by using

$$\int_0^L dx \frac{\partial p}{\partial x} = p_L - p_0 = 0 \tag{7.28}$$

which gives

$$Q = V\frac{h_0 h_L}{h_0 + h_L} \tag{7.29}$$

We now integrate the differential equation directly:

$$p(x) = p_a + \frac{6\eta V\alpha}{h_0 + h_L}\frac{x(L-x)}{(h_0 - \alpha x)^2} \tag{7.30}$$

Thus pressure is a parabolic function, with a maximum between $x = 0$ and $x = L$. This leads to a lift force on the upper body, which in principle can balance its body weight:

$$F_L = L_y \int_0^L p dx = p_a L_y L + \frac{6\eta V L^2 L_y}{h_0^2} G(k) \tag{7.31}$$

with the geometrical parameter $k = h_0/\alpha L$, $1 \leq k \leq \infty$, and

$$G(k) = k^2(\ln \frac{k}{k-1} - \frac{2}{2k-1}) \tag{7.32}$$

which diverges for $k \to 1$ and decays to smaller values with increasing k. Note that if the lower surface moves in the opposite direction, the lubrication force will pull the object down.

7.3.2 Viscous adhesion

An important application of the lubrication approximation is viscous adhesion, eg between a surface and an adhesive tape. Consider two disks of radius R each which approach each other with velocity V, thereby squeezing out fluid to the sides. We use cylindrical coordinates and again use the approximation of a thin film. Then the Stokes equation gives

$$\eta \frac{\partial^2 v_r}{\partial z^2} = \frac{\partial p}{\partial r} \tag{7.33}$$

and the equation of continuity reads

$$\frac{1}{r}\frac{\partial (r v_r)}{\partial r} + \frac{\partial v_z}{\partial z} = 0 \tag{7.34}$$

The boundary conditions are $v_r = v_z = 0$ at $z = 0$, $v_r = 0, v_z = -V$ at $z = h$ and $p = p_0$ at $r = R$. Integrating the Stokes equation gives

$$v_r = \frac{1}{2\eta}\frac{\partial p}{\partial r} z(z-h) \tag{7.35}$$

Using the continuity equation gives

$$\int_0^h dz \frac{\partial v_z}{\partial z} = V = -\frac{1}{r}\frac{\partial}{\partial r} r \int_0^h dz v_r = \frac{-h^3}{12\eta r}\frac{\partial}{\partial r}(r\frac{\partial p}{\partial r}) \tag{7.36}$$

This can be integrated to give

$$p = p_0 + \frac{3\eta V}{h^3}(R^2 - r^2) \tag{7.37}$$

The increased pressure inside leads to a viscous force resisting the imposed movement:

$$F = 2\pi \int_0^R r dr p = \pi R^2 p_0 + \frac{3\pi\eta V R^4}{2h^3} \tag{7.38}$$

Note the strong scaling with h. This explains why it is much easier to peel an adhesive tape rather than to pull it off vertically.

7.3.3 Stokes drag for a sphere

The translating sphere in a resting fluid is equivalent to the resting sphere in a moving fluid. Consider the fluid moving along the negative z-direction with velocity v. This problem was solved by Stokes in 1851:

$$v_r = v\cos\theta\left(1 - \frac{3R}{2r} + \frac{R^3}{2r^3}\right) \tag{7.39}$$

$$v_\theta = -v\sin\theta\left(1 - \frac{3R}{4r} - \frac{R^3}{4r^3}\right) \tag{7.40}$$

$$p = p_0 - \frac{3\eta R}{2}\frac{v\cos\theta}{r^2} \tag{7.41}$$

We now calculate the Stokes force

$$F_S = \int dA(-p\cos\theta + \sigma_{rr}\cos\theta - \sigma_{r\theta}\sin\theta) \tag{7.42}$$

where the integral is over the sphere surface and the force is projected onto the z-direction. The stress tensor can be calculated from the velocity field. We specifiy it together with the pressure for the sphere surface:

$$\sigma_{rr} = 0,\ \sigma_{r\theta} = -\frac{3\eta}{2R}v\sin\theta,\ p = p_0 - \frac{3\eta}{2R}v\cos\theta \tag{7.43}$$

Then we get

$$F_S = \frac{3\eta v}{2R}\int dA = 6\pi\eta R v \tag{7.44}$$

Note that the scaling is unique and therefore is also obeyed by non-spherical bodies; then R is the dominant linear extension, for example the long axis of an ellipsoid. Also note that the Stokes solution is only valid in the limit $Re = 0$. For finite Reynolds number it is not valid in the far field, and one can calculate corrections as an expansion in small Reynolds number.

We briefly comment on the relation between the Stokes drag and the diffusion constant of colloids. From the Stokes drag, we get a friction coefficient $\Gamma = F/v = 6\pi\eta R$. Its inverse is the mobility, $M = 1/\Gamma = 1/6\pi\eta R$, because $v = MF$. Einstein proved that this dynamic response should also show up in equilibrium, namely as a diffusion constant (this is a special case of the fluctuation-dissipation theorem, which connects the dissipative response to

an external force with the fluctuations in equilibrium). We start by noting that in equilibrium the overall flux has to vanish:

$$j = -D\nabla c + vc = 0 \tag{7.45}$$

where the first and second term represent diffusion and convection, respectively. For velocity we have $v = MF = -M\nabla U$ with some potential U. Therefore we get

$$D\nabla c = -cM\nabla U \;\; \Rightarrow c = e^{-MU/D} \tag{7.46}$$

Because the equilibrium distribution should at the same time be the Boltzmann distribution e^{-U/k_BT}, we get the famous *Einstein relation*:

$$D = Mk_BT = \frac{k_BT}{6\pi\eta R} \tag{7.47}$$

In anisotropic situations, e.g. for a sphere in front of a wall, mobility becomes direction-dependent, that is scalar mobility is replaced by a mobility matrix. In the general case one also has rotational degrees of freedom, which can be treated in the same hydrodynamic framework.