Lecture notes: **Physical Chemistry of Interfaces and Colloids**

Diethelm Johannsmann, Institute of Physical Chemistry, TUC

These notes contain background information. When preparing for the exam, also look at the questions from previous exams.
Last updated: 11/29/2024

Content

1 General, literature

In the introduction to Physical Chemistry, "phases" play a prominent role. There is talk of "almost all" molecules and "almost all" means: all molecules except those at the interfaces. The term "almost all" becomes problematic when the size of the domains becomes small and the surface-to-volume ratio becomes large. This is the case with colloids. The course does, however, not start with the colloids, but with the interfaces themselves. Interfaces may have some finite thickness, but they are not phases. Often they are molecularly sharp and molecular details therefore are important. Interfaces often carry function, not only in colloids.

Good books are:

- Butt, H.-J. et al. *Physics and Chemistry of Interfaces* [1]
- Goodwin, J., *Colloids and Interfaces with Surfactants and Polymers* [2].
- Morrison, I. D et al, *Colloidal Dispersions* [3].

These are lecture notes, this is not a textbook. Because lecture notes change, inaccuracies and errors occasionally creep in. If you find room for improvement, please send an e-mail to johannsmann@pc.tu-clausthal.de.

I spent time to diligently formulate the German version of this text. The translation was done with DeepL and I spent limited time to catch errors. Some figures still contain German labels.

2 The surface energy

2.1 General

Interfaces are characteristic of *phase coexistence*. A phase is a region in which the materials parameters are either constant or vary slowly. The latter may be caused, for example, by gravitation (as in the atmosphere).[4] At interfaces between phases, the physical parameters change discontinuously (or at least almost so). Because the jump is sharp, the structure and the behavior of interfaces is influenced by details of the molecular organization. For example, most ions dissolved in water increase the surface energy. Some types of ions, however, decrease it.

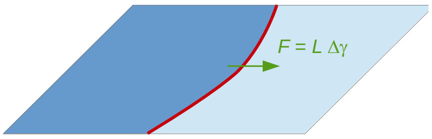


Fig. 2.1
If a water surface is divided into two compartments with a thread and when a surfactant is spread onto one of the compartments, the surfactant exerts a force on the thread. The force is caused by the difference in the surface energies. It is proportional to the length of the filament. This illustrates that γ can also be viewed as a force per length.

At interfaces, the free energy per molecule is higher than in the bulk. If this increased energy (more precisely: the increased free enthalpy) is normalized to the surface area, one arrives at the surface energy, γ . As long as γ is positive, the system tries to minimize its surface. Drops and bubbles therefore tend toward spherical shape. Negative surface energy, which occurs in connection with electrocapillarity and which leads to an instability, will be discussed in Sect. 2.4.

The dimension of the surface energy is either J/m^2 or N/m. (Remember: 1 J = 1 Nm.) The second view is close to intuition. Fig. 2.1 shows a planar water surface, which is divided into two compartments by a thread. Let the surface on the left be covered by a molecular monolayer of a surfactant, which is insoluble in water. To be precise, let the solubility of

---

[1] Butt, H.-J.; Graf, K.; Kappl, M., *Physics and Chemistry of Interfaces*. Wiley-VCH: 2006
[2] Goodwin, J., *Colloids and Interfaces with Surfactants and Polymers*. Wiley: 2004
[3] Morrison, I. D.; Ross, S., *Colloidal Dispersions*. Wiley: 2002
[4] For nematic liquid crystals, the preferred direction can vary slowly in space.

the surfactant be so low that the number of molecules on the surface can be considered constant. If the surfactant is largely insoluble, it cannot migrate through the bulk phase to the other side of the thread. The surfactant lowers the surface tension. The difference in surface tension Δγ between left and right results in a force onto the filament, which is proportional to its length. Δγ is the same as the force divided by the length of the filament. This force-length ratio is the 2D analog of the pressure in 3D.[5] This analogy will be discussed in more detail to in Sect. 7.2. The 2D pressure exerted by surfactants causes surfactants to spread spontaneously on a water surface. If a hydrophobic wire is bent into a horseshoe, if this horseshoe is carefully placed on the water surface so that it does not sink, and if a surfactant is sprayed into the space between the two arms, the recoil of the surfactant flowing out of the horseshoe at the open end propels the horseshoe forward.

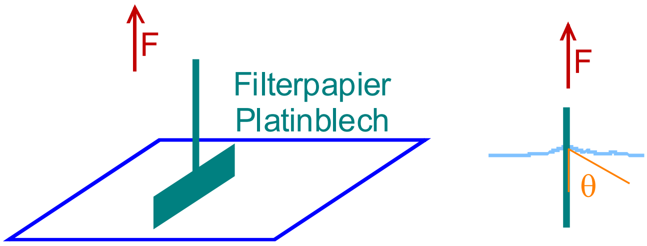


Fig. 2.2
When a platinum sheet (or a filter paper) is immersed in a liquid, the surface of the liquid exerts a force proportional to the length of the three-phase line and the surface energy γ. γ can be inferred from that force. The platinum sheet or the filter paper is also called "Wilhelmy plate". Since the contact angle and the exact circumference of the Wilhelmy plate at the meniscus are not known, calibration is usually carried out with pure water.

Fig. 2.2 shows a second example, which is the Wilhelmy plate. The plate (it can also be a piece of filter paper) is immersed in the liquid. The meniscus exerts a force onto the plate. This force is proportional to the surface tension and the length of the three-phase line.

Guided by intuition, one might be tempted to identify surface tension with capillary pressure (section 3). Surface tension and capillary pressure are related, but they are not the same. The units are different (N/m and N/m^2).

A note on terminology: Surface tension is largely equivalent to surface energy. Keep track of the units: surface tension as units of N/m, whereas the conventional tension has units N/m^2. Surface tension can be viewed as a tension in 2D, hence the unit N/m instead of N/m^2.

It is common to report surface energies in units of mN/m (mN for milliNewton). The prefix "milli" achieves two things. First, the numerical values are the same as in the cgs system (in the centimeter/gram/second system instead of the meter/kilogram/second system, the latter also called the "SI" system[6]). In the cgs system, the unit for force is the dyne. The analog of the Joule is the erg (1 erg = 10^{-7} J). The dyne (abbreviated as dyn) and the Newton are related by 1 dyn = 10^{-5} N. The quantity mN/m corresponds to one dyn/cm. Second, the unit mN/m makes it easy to pronounce the numerical values of γ for the ordinary liquids. For example, water has a surface energy 72 mN/m (same as 0.072 N/m, but the latter is less easily pronounced).

Table 1.1 Values of the surface energy for selected substances

	γ [mN/m] (20°C)
Helium (2–4 K)	0.12 – 0.35
n-Pentan	16
PTFE (Teflon)	18
Silicone	20 – 25
Benzene	29
Water	73
Mercury	486
Aluminum	1100
Tungsten	4400

---

[5] The pressure in 3D has either the unit J/m^3 or the unit N/m^2 .

[6] The cgs system is common in North America, the SI system in Europe. The cgs system has nothing to do with the inches, pounds, and miles-per-gallon that US-Americans use to celebrate their cultural roots. It is a "metric" system.

Some values should be remembered. Among the different liquids, water has a relatively high surface energy (γ = 72 mN/m) because the internal cohesion is high due to the H-bonds. Mercury has an even higher surface energy (486 mN/m) because of the metallic bonds. Solids also have a surface energy. However, it is not easily measured because solids do not form droplets. Still, the surface energy as such does exist and it actually is greater than the surface energy of most liquids[7] because of the high internal cohesion. Tungsten (one of the "refractory metals") has both a high surface energy and a high melting point.

Low surface energies are of some practical importance. At room temperature, the surface energies of ordinary liquids can hardly fall below about 10 mN/m, because the material would then have such low internal cohesion that it would evaporate.

Fig. 2.3 shows the surface energies of liquid noble gases, plotted against the atomic number. This is a homologous series, i.e. a series of substances that are similar to each other except for a single parameter (here: the atomic number). As can be seen, the surface energy is correlated with the vaporization enthalpy (Sect. 2.3), with the boiling temperature, with the refractive index, and with the polarizability. This is an example, for which the relationship between the properties of molecules (in this case, atoms) and the properties of the corresponding macroscopic substances is well understood. The larger the atom, the larger the number of electrons and the larger the electronic polarizability. In turn, the electronic polarizability[8] is correlated with the refractive index (an optical property). One can convert electronic polarizability and refractive index into each other using the Lorentz-Lorenz equation. A high polarizability also entails large quantum mechanical fluctuations of the dipole moment. The correlated fluctuations of the dipole moment generate the London dispersion interaction (one of the three contributions to van-der-Waals attraction). Therefore, as the atomic number increases, so does the attraction between neighboring atoms. As a consequence, the evaporation enthalpy, the boiling point and the surface tension all increase.

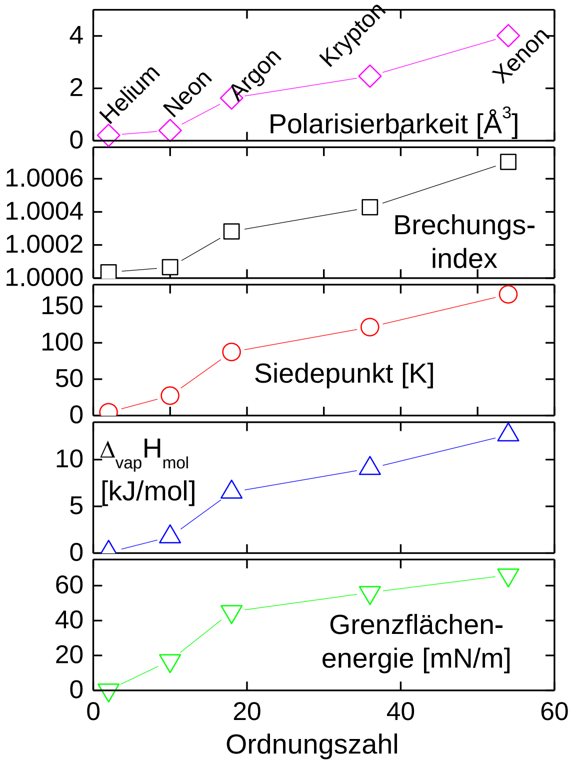


Fig. 2.3
An example of the relationship between the properties of certain atoms on the one hand and the macroscopic properties of the phases concerned on the other. This relationship is also called "structure-property relationship". The substances are a "homologous series".

Remember

- Interfaces usually carry a (free) energy. In liquid phases, the system minimizes its surface area, which leads to a spherical shape of drops or bubbles.
- The unit of surface energy is mN/m (milliN/m). Typical values are a few 10's of mN/m. Water at 25°C has a surface energy of γ = 72 mN/m.

---

[7] Polymers are considered liquids in this context because the bonds exist only along the chain.

[8] There also is a polarizability, which results from the displacement of whole atoms with respect to each other, and, furthermore, a polarization, which results from a preferential orientation of molecules with permanent dipole moments in an external field. These two effects lead to "atomic polarization" and "orientation polarization". Both mechanisms are too slow to follow the frequency of light ($\approx 10^{15}$ Hz) and therefore have no influence on the refractive index.

– The surface energy is high for materials with high internal cohesion.

2.2 Thermodynamic treatment of the surface energy

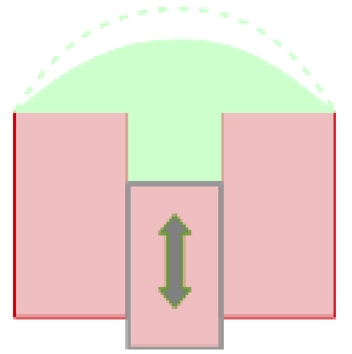


Fig. 2.4
When the piston is moved, the interface between the liquid and the vapor is changed by an amount dA. This change in area associated with a surface work $dG_{surf} = \gamma dA$.

The thermodynamic treatment of interfaces contains a few subtleties, which we avoid for now by making assumptions. To the extent that we drop these assumptions later, it will become clear how complicated interfaces can be. Consider the surface shown at the top in Fig. 2.4. A liquid is in contact with a vapor phase. By moving the piston up, the area of the interface between the liquid and a vapor is increased. In the process, a surface work is done. The assumptions are:

a) Liquid and vapor are in thermodynamic equilibrium.
b) The interface is almost molecularly sharp.[9] Because it is sharp, it is easy to see where, exactly, it lies. Even for strongly curved interfaces (such as the surfaces of nanodroplets), the surface area (given as $A = 4\pi r^2$ for spherical droplets) can be specified with good accuracy.[10]
c) No work other than the surface work is done when the piston moves. The piston moves without friction. All other surface energies in the system are zero. Gravity is negligible. The pressure-volume work on the vapor phase is compensated by an equal pressure-volume work on the back of the piston. The piston and the tube are rigid bodies, meaning that no elastic work is done.
d) There are no impurities.
e) The three-phase boundary between the tube, the liquid, and the vapor is anchored at the outer edge of the tube. The meniscus is immobile ("pinned"). Only the area of interface between the liquid and the vapor changes.
f) There is no evaporation or condensation while the piston moves.

---

[9] The exceptions are worth a mention:
- Interfaces between liquid and vapor near the critical point are not molecularly sharp. In two-phase systems close to the critical point (the critical point of segregation or the critical point of vaporization), both phases have almost equal properties. Then the surface energy tends to zero; the surface width tends to infinity. Exactly at the critical point, the interface finally disappears completely.
- Interfaces of electrolytes are still not molecularly sharp (keyword: diffuse double layer, section 10.5).
- Interfaces between moderately incompatible polymers are still not molecularly sharp. Here, a molecularly sharp interface would force a refolding of chains at the interface and thus disturb the tangle statistics. The resulting small widening of the junction is of practical importance because it can lead to hooking between the different polymer materials ("polymer welding"). Entanglements are favorable for adhesion. Note that the debonding work per area is always greater than the equilibrium surface energy. In the favorable cases, it is greater by a factor of 1000 or more. Adhesives cause work to be done during delamination that far exceeds the equilibrium surface energy. This can include the unwinding of polymer chains. Most of this energy is dissipated as heat. It may also remain in the system as chemical energy (chain fracture). In the latter two cases, the mixed region is sometimes called "interphase" instead of "interface." This term is problematic because the interface – even if it is broad – is not a phase in the thermodynamic sense. An exception is the solid-electrolyte interphase (SEI) in lithium-ion batteries. This *is* a phase. It consists of decomposition products of the electrolyte. It may be thought of as a film.

[10] Because interfaces between phases are molecularly sharp, the molecular details matter. This coin has two sides. On the one hand, it makes the description more complicated. On the other hand, molecular interactions and molecular processes can be studied at interfaces with moderate effort. An example: Ions like to stay in hydrophilic environments, because the hydration shell and the counterion cloud lower the energy. Therefore, ions are usually *depleted* at the water-air interface. As discussed in Sect. 5.25.2, depletion increases the surface tension. Based on a measurement of the surface tension, one may quantify depletion (equal to the negative Γ_2). It turns out that some ions actually *lower* the surface tension. The argument "hydration shell and counter ion cloud → depletion " does not always apply. The reasons for these peculiar "ion-specific effects" are not quickly explained or even well understood. However, an enrichment of ions on hydrophobic surfaces is of outstanding importance for molecular biophysics. The water/air interface is a model system for the study of ion-specific effects.

For this situation and under these conditions, the surface energy, γ, is defined as

$$\gamma = \left(\frac{\mathrm{d}G}{\mathrm{d}A} \right)_{p,T} \qquad \text{Eq. 2.2.1}$$

G is the Gibbs Free Enthalpy of the entire system and A is the area of the liquid-vapor interface. When using the Gibbs Free Enthalpy, it is implied that all derivatives are to be carried out at constant pressure and temperature. Recall that the Gibbs Free Enthalpy is given as $G = U + pV - TS$ and that systems under isothermal and isobaric boundary conditions tend towards minimum G. At constant temperature and pressure, minimal G defines thermodynamic equilibrium. *In assumption a) we restricted the scope to systems in thermodynamic equilibrium*. These are the liquids and gases (often collectively referred to as "fluids"). For fluids, G is usually is equal to $G_{eq} = G_{min}$. Note: This restriction is not strictly needed for a discussion of the surface energy. Solids (which are not usually in thermodynamic equilibrium) also have a surface energy, see sect. 2.5).

The local thermodynamic equilibrium is established at interfaces between fluids over a time of a few nanoseconds. That is the residence times of individual molecules at the interface. After this time, they are exchanged for other molecules from the bulk of the adjacent phases.

The area A is well-defined because phase boundaries are molecularly sharp (assumption *b*). If the time-averaged density is plotted versus z (z the normal coordinate at the interface, cf. Fig. 2.5), this density drops continuously at the interface, but the drop is steep.[11] The energetic conditions are more favorable in the bulk phases than in the transition region, and the system therefore keeps the transition region as narrow as possible. For the exceptions, see footnote 10.

In the formulation of it was assumed that the piston only does work on the interface. No work is done on the bulk phases (assumption *c*). To be precise, in the experiment sketched Fig. 2.4 the work against gravity (which is nonzero) has to be subtracted from the total work to obtain γ. A similar problem arises (in a form not easily dealt with) when measuring the surface energy of solids (Sect. 2.5). One cannot change the interface of a solid without doing elastic work at the same time. This elastic work must be subtracted from the total work.[12] A related topic: When droplets rest on deformable bodies (e.g. on elastomers or on gels), the capillary pressure inside the drops deforms the substrate, which leads to "elastocapillarity". Elastocapillarity is of much importance when the capillary pressure is high. This is the case in many porous and granular media. Another form of non-surface work is pressure-volume work on air bubbles. This work is important for large air bubbles (with small capillary pressure, Sect. 3). It is insignificant for liquid drops because these are almost incompressible.

---

[11] There is a complication here, which goes back to *thermal capillary waves.* Capillary waves (more generally: waves on the surfaces of liquids) are complicated and we will not go into details. At finite temperature, every fluid interface exhibits dynamic ripples, even without being excited in any way. These are a form of Brownian motion. The ripples contribute to the broadening of the density profile. This has two consequences:

- When averaging over large areas to calculate the density profile, the low Fourier components contribute to the width of the transition. Otherwise they do not. Therefore, the shape of the profile $\rho(z)$ depends on the area of averaging. (A related problem occurs for roughness. The roughness value depends on the area of averaging.)
- One might be tempted to ignore the capillary waves in some way and calculate a local, "intrinsic" interface width. However, there is no clear prescription of how to distinguish between the motion of single molecules and small-scale capillary waves. The intrinsic width of the interface depends on what is meant by "local".

[12] If interested in the details, search the web for the Shuttleworth equation.

We avoided transport equilibria in the context of by considering only one species (assumption *d*). Transport equilibria at interfaces are important in practice because interfaces are rarely clean. Note that adsorption equilibria often develop slowly because diffusion takes time because adsorbate layers often reorganize internally.

The anchored contact line (assumption *e*) is only necessary if one wants to write down simple relations between the movement of the piston and the change of *A*. The definition of the surface energy per se is independent of the anchoring the meniscus. Of course, one has to take the surface energy between solid substrate and fluid into account in case the contact line slides on the substrate (Sect. 6). An anchoring of the meniscus (as in Fig. 2.4) is quite realistic.

Evaporation and condensation must be excluded because work is associated with it (Sect. 4.1).

So much for the definition of surface energy. We now turn to a conceptual subtlety. One can assign a Gibbs free enthalpy to the interface according to

Eq. 2.2.2

$$G_\sigma = G_{tot} - G_\alpha - G_\beta$$

The indices α and β denote the two bulk phases here. The index σ denotes the interface. One has

Eq. 2.2.3

$$G_\sigma = A\gamma$$

with γ from . G_σ is an extensive quantity. *However, the Free Enthalpy associated with the interface does not imply that the interface itself was a "phase".* This latter notion underlies the Guggenheim picture (blue area at the top of Fig. 2.5). According to Guggenheim, the interface would correspond to a certain spatial region of approximately the width of the transition in Fig. 2.5. However, the Guggenheim picture is misleading. This can be seen from the fact that it is not clear how wide the strip in question would have to be. But there is a more fundamental problem: The Gibbs free enthalpy *G* has no well-defined zero. *Differences* in *G* drive spontaneous processes; *G* is not defined absolutely. The surface energy, however, can be defined and measured in absolute terms. It can also become negative (Sect. 2.4). This cannot be understood within the framework of the Guggenheim picture.

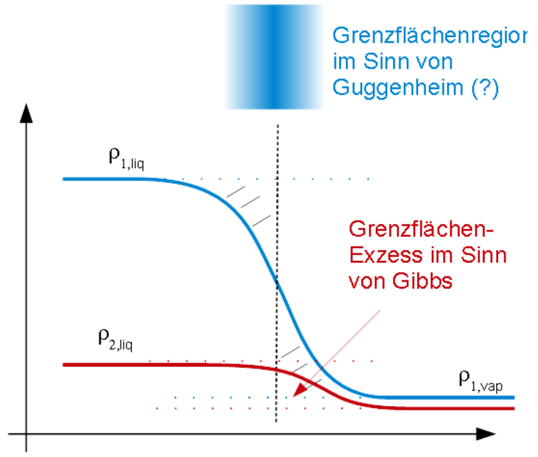


Fig. 2.5
The Gibbs dividing surface (GDS) is chosen such that the surface excess of the majority component is zero. Values of the surface excess of the other components then are nonzero.

Differing from Guggenheim picture, G_σ is an *excess quantity*. For the definition of excess quantities, we consider only cases where there is a majority component. A majority component is either a pure substance or a solvent in which the solute is diluted. By convention, the majority component is given the index 1. Near the surface, the time-averaged density $\rho_1(z)$ decreases continuously. Again: This continuum picture neglects the molecular structure. Fluid-fluid interfaces are almost molecularly sharp and the molecular details do matter. Nevertheless, one can always determine a time-averaged density $\rho_1(z)$. This density will gradually decrease at the interface. In the next step, one compares the profile $\rho_1(z)$ with a hypothetical profile with an infinitely sharp interface (dashed vertical line in Fig. 2.5). This hypothetical interface is the "Gibbs dividing surface" (GDS). The exact position of this interface on the z-axis is fixed by the requirement that the integral over the difference between real density profile $\rho_1(z)$ and hypothetical density profile with a sharp edge is zero. The two dashed areas in Fig. 2.5 must be equal in size. The integral over this difference is also called the "interface excess". The interface excess of the density of the majority component (only of the majority component) is always zero according to the definition of Gibbs dividing surface.[13]

A surface excess can be calculated for any other density. For all densities except ρ_1, the surface excess is non-zero in general. This includes the density of enthalpy, entropy, free energy, and the density of solutes. One always extrapolates the bulk values of the density on both sides to the GDS and subtracts this box profile from the real profile. This extrapolation procedure ensures that all excess quantities have a well-defined zero. If the density under consideration is the density of a substance in units of mol/L, it may be enriched at the interface (as is the case for many surfactants in water) or depleted (as is the case for many ions in water, Sect. 5).

By means of the surface excess, extensive quantities are defined for interfaces and one can apply the formalism of thermodynamics to interfaces. For example, one can write

Eq. 2.2.4

$$\mathrm{d}G_\sigma = \gamma \mathrm{d}A = \mathrm{d}U_\sigma + V_\sigma \mathrm{d}p - S_\sigma \mathrm{d}T$$

Two complications: From the equation above, we derive a "surface volume" and a "surface entropy". The "surface volume" is defined as

Eq. 2.2.5

$$\frac{V_\sigma}{A} = \left(\frac{\mathrm{d}\gamma}{\mathrm{d}p}\right)_{T,n} = \left(\frac{\mathrm{d}^2 G}{\mathrm{d}A\mathrm{d}p}\right)_{T,n}$$

If one naively defines a "volume density" (i.e. volume/volume), then this density is trivially equal to one everywhere and there can be no excess volume. But one can also define the volume density as the volume per particle. Then the formalism is consistent again.

The surface entropy is defined as

Eq. 2.2.6

$$\frac{S_\sigma}{A} = -\left(\frac{\mathrm{d}\gamma}{\mathrm{d}T}\right)_{p} = -\left(\frac{\mathrm{d}^2 G}{\mathrm{d}A\mathrm{d}T}\right)_{p}$$

A word of caution: It is problematic to define an "entropy density" as an average value over a volume corresponding to a molecular volume. Entropy is related to the number of microconformations of a "system". For systems where there are interactions on the one hand and steep density gradients on the other, it is not possible to define an entropy density as a function of z at the molecular volume level. The simplest way is to start again directly from the derivation of γ according to the temperature.[14]

---

[13] If a majority component does not exist, the determination of the location of the GDS becomes more complicated. One must then use that the interface excess of the G-density is equal to γ according to . We do not prove that this is also the case for the definition of the GDS using the majority component.

[14] In aqueous systems, the analysis of the temperature dependence of γ has its limitations because many bulk properties of electrolytes are also temperature-dependent.

As a rule, the surface energy decreases with temperature. From this, one concludes that the surface entropy (an excess entropy) is positive. This is not trivial, but nevertheless plausible. At the interface between two phases, it is easy to imagine that the number of possible microstates per volume is higher than in the homogeneous adjacent bulk phases. At this point, thermal capillary waves also play a role (footnote 11).

Note: If the two bulk values to the right and left of the interface are different, the value of the relevant excess size depends on the location of the GDS. In Fig. 2.5, we had placed the GDS in such a way that the interface excess of the majority component is zero. Of course, this procedure is viable only if there is such a majority component. The general rule is: the GDS must be such *that* the excess quantity in question (call it B^σ) corresponds to the derivative of the quantity in question with respect to the surface ($B^\sigma = \mathrm{d}B/\mathrm{d}A$). This is the way to proceed; the concept is consistent.

One last remark: Be careful not to assign a chemical potential to the molecules at interfaces in one-component systems. To do this, one would have to write

Eq. 2.2.7

$$\mathrm{d}G_\sigma = \gamma \mathrm{d}A + \mu_1 \mathrm{d}n_{\sigma,1}$$

However, this notation is misleading because the area A and the amount of substance at the interface $n_{\sigma,1}$ cannot be varied independently. This is another consequence of the fact that the interface is not a phase. For multicomponent systems, one can define chemical potentials at interfaces (Sect. 5) but the number of these well-defined chemical potentials is always 1 less than the number of chemical components.

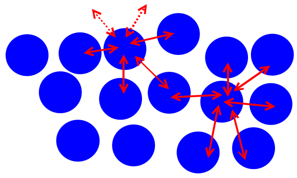


Fig. 2.6
In simple models, the surface energy arises from the fact that molecules at the surface have fewer neighbors than molecules in the bulk.

One may question the practical benefit obtained from the Gibbs formalism, if it is so complicated. For the practice it is in fact mostly more purposeful to start directly from the derivation of the quantity in question according to the area (as in Eq. 2.2.5). The notion of an excess quantity is a conceptual bridge by means of which one can see that the same thermodynamic principles apply at interfaces and in the bulk. The Gibbs formalism primarily is applied the context of the Gibbs adsorption isotherm (Sect. 5.2). There, the excess refers to a quantity of substance.

Remember

- The surface energy is defined as $\gamma = (\mathrm{d}G/\mathrm{d}A)_{p,T}$.
- The interface is not a phase of its own.

2.3 The surface energy as a consequence of the missing neighbors

The processes and conditions at surfaces can be extraordinarily complicated. This includes the surface energy. In particular, there may be *interface anomalies*. A material can behave differently at an interface than it does in the bulk. An important example is water, which often orders at interfaces.[15] However, that is not to say the surface energy is always and without exception linked to an anomaly. One might suspect some kind of skin at interfaces which forces the droplet into a spherical shape, just

---

[15] Sometimes one speaks of "ice-like" structures. This term is problematic in that although the order at the interface of liquid water is high, the structure is not the same as the structure of ice.

as a balloon forces the gas volume inside into an almost spherical shape. However, the surface energy can be explained without recourse to a skin. It can be caused by missing neighbors alone (Fig. 2.6).

Let ε be the interaction energy between two molecules, let *z be* the average number of neighbors of a molecule (the "coordination number"), and let A_m be the area per molecule. The surface energy can then be estimated as

Eq. 2.3.1

$$\gamma \approx \frac{1}{A_m}\frac{z}{2}\frac{\varepsilon}{2}$$

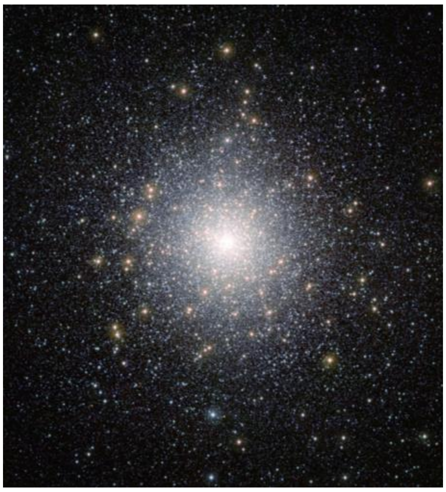


Fig. 2.7
47 Tucanae, the second brightest globular cluster in the Milky Way.
Globular clusters consist of several 100 000 stars, which are loosely bound to each other by gravity. Since gravity is a long-range interaction, there is no sharp interface between the interior of the cluster and the surrounding galactic space. Such a sharp interface would require a short-range interaction.
en.wikipedia.org/wiki/Bullet Cluster

The factor 1/2 at *z* occurs because about half of the neighbors are missing. The factor 1/2 at ε occurs because the interaction always links two partners and must not be counted twice. This simple equation has a number of implications:

- The surface energy should be correlated with the evaporation energy. We follow an elementary argument of Young from 1815. Young derives an estimate for size of the molecules d_M from the ratio of $\Delta_{vap}\bar{U}$ and γ . Young glosses over complications arising from the coordination number. He argued with the dimensions: The size $\Delta_{vap}\bar{U}/\bar{V}$ is an energy per volume γ is an energy per area. Then the ratio/($\Delta_{vap}\bar{U}/\bar{V}$) has the dimension of a length. This length should be approximately the molecule diameter. We use the values for water ($\gamma = 72{\cdot}10^{-3}$ J/m^2, $\Delta_{vap}\bar{H} = 44$ kJ/mol, $\Delta_{vap}\bar{U} = \Delta_{vap}\bar{H} - RT$, $\bar{V} = 18\ 10^{-6}$ m^3) and obtain

  Eq. 2.3.2

  $$d_M \approx \frac{\gamma\bar{V}}{\Delta_{vap}\bar{H} - RT} = 0.03\ \text{nm}$$

  This is about a factor of 10 below the experimental value. The agreement can be improved by various refinements (e.g., a consideration of the coordination number), but we leave the argument as it is.

- The surface energy is correlated with the internal cohesion of the material. Mercury, for example, has a high surface energy (486 mN/m) due to the metallic bonds. Water also has a fairly high surface energy (72 mN/m) because of the H-bonds. The surface energy of benzene – mostly held together by van-der-Waals forces – is 30 mN/m. The surface energy of fluoropolymers (e.g. Teflon) is particularly low because the fluorine atom is poorly polarizable.[16]

- A final remark on the missing-neighbor model concerns not its consequences, but the assumptions entering it. The model assumes that a given molecule interacts primarily with its *neighbors* – possibly also with the second shell of neighbors. The model presupposes a *short-range* interaction. The van-der-Waals attraction ($\propto r^{-6}$) is a short-range interaction in this sense.

---

[16] Teflon is rather expensive. If cost is an issue, silicones are an option. Why silicones have low surface energy is not easily explained. Silicones contain segments of the form -O-(SiR_1R_2)- in the main chain

On the subject of short-range interaction, there is an interesting historical note from nuclear physics. In the early days of nuclear physics, it was shown that there is a contribution to the energy of the nucleus that scales like $N^{2/3}$, where N is the number of nucleons (protons and neutrons).[17] For nuclei with spherical shape, the surface area also scales like $N^{2/3}$ (like $V^{2/3}$ with V the volume, the latter proportional to N) and it is natural to interpret the $N^{2/3}$-term as a surface energy. This led to the "droplet model" of the nucleus. This model makes only sense if the interaction between the nucleons (the nuclear force, also: the "strong interaction") is indeed short-ranged. It is; as was confirmed by scattering experiments.[18] The surface term in the nuclear energy explains the energy gain in nuclear fusion. The energy gain in nuclear fission has a different cause. It results from the competition between short-range attraction (caused by the nuclear force) and long-range repulsion (caused by the electric repulsion between the protons). We will come back to such a competition in Sect. 7.1.

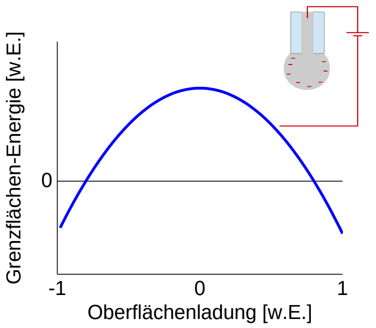


Fig. 2.8
When you charge a conductive droplet, the charge accumulates on the surface. Then the surface tension decreases in proportion to the square of the charge. The decreased surface tension is caused by to the repulsion between charges of the same sign.

There *are* aggregates which are held together by long-range attractive interactions, but which do not have a sharp interface because of the long range attraction. These are the globular clusters. Fig. 2.7 shows a picture. Globular clusters contain some 100 000 stars, which move (on the time scale of many millions of years) randomly. A statistical analysis is appropriate and successful. For stars, interaction with distant partners is essential. While the interaction potential scales like r^{-1} , the number of stars in a shell of radius r scales like r^2.[19] A star interacts with all other stars (neighbors or not). The density profile $\rho(r)$ (ρ the number density of stars) therefore decreases continuously at the edge of the cluster; there is no sharp edge.

---

[17] This binding energy is determined by mass spectroscopy using the relation $E = mc^2$. Because of the binding energy, the mass of a nucleus is not exactly equal to the sum of the masses of the nucleons (the protons and the neutrons). The difference is called the "mass defect". One can determine the sum formula of a molecule with high-resolution mass spectroscopy. (For example, one can distinguish 12 H atoms from one C atom.)
In principle, this procedure might also be applied to the energy of chemical bonds. However, this would require to determine the molecular mass to within 10 decimal places, since 1 eV is roughly equivalent to 10^{-9} atomic mass units. The current accuracy of mass spectrometers is 4 – 5 decimal places. This suffices to determine the binding energies in the nucleus (in the range of mega-electron-volts, MeVs), but not the binding energies in molecules (a few eV).

[18] When two particles with short-range interaction almost collide, they do not change direction because they do not see each other. For the Coulomb interaction, there is a small change in direction even if the particles do not come close.

[19] Assuming constant density, which is questionable.

Because molecules interact primarily with their nearest neighbors, liquid/liquid interfaces and liquid/vapor interfaces are almost atomically sharp. The fact that interfaces are sharp implies that there is a clear difference between "surface" and "bulk" in nanoparticles. Extremely small aggregates, where one can no longer make this distinction, are called "clusters".

Remember

- Although there are many types of surface anomalies, the surface tension is not necessarily associated with a "skin". The missing neighbors alone also create a surface energy.
- The missing-neighbor model requires short-range interactions (interactions with the neighbors only).
- The missing-neighbor model explains why surface energy is correlated with internal cohesion.

2.4 Electrocapillarity and negative surface energies

A negative surface energy cannot be a property of an equilibrium state, because such an interface will always strive to increase in size. At the end of such a process, the bulk phase in question will dissolve completely.[20]

Consider a conductive liquid (e.g. mercury or any electrolyte), which is electrically charged by means of an inserted electrode. Since all electric fields inside the drop are dissipated by corresponding electric currents, and since, according to Maxwell's 1st law, the charge density is equal to the divergence of the field ($\rho = \nabla \cdot E$), the relation $\rho = 0$ follows from $E = 0$. The entire charge is therefore located at the surface. Since charges of the same sign repel each other, the charge at the surface *lowers the* surface energy. The Lippmann equation states that:

Eq. 2.4.1

$$\sigma_M = -\frac{d\gamma}{d\Delta\varphi}$$

σ_M is the surface charge density, $\Delta\varphi$ is the potential jump at the interface. (Eq. 2.4.1 can be made plausible by the relation $\delta\gamma = -\sigma_M\, d\Delta\varphi$). If $\sigma_M = C\,\Delta\varphi$ with C the capacitance, integration leads to

$$\gamma \approx \gamma_0 - \frac{1}{2} C \left(\Delta\varphi - \Delta\varphi_0\right)^2$$

Eq. 2.4.2

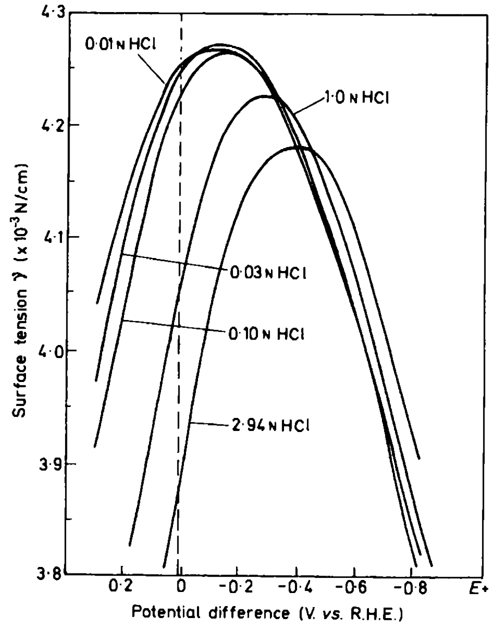


Fig. 20.6 Electrocapillary curves for HCl at various concentrations determined using a reversible hydrogen electrode (R.H.E.) immersed in the same concentration of HCl as that used for the determination (after Bockris and Reddy[1])

Fig. 2.9
The surface tension of mercury in different salt solutions
chempedia.info/info/4739/

---

[20] The lamellar phases, that sometimes result from self-assembly of surfactants, can be viewed as the result of such a process. We will come back to self-organization in Sect. 7.

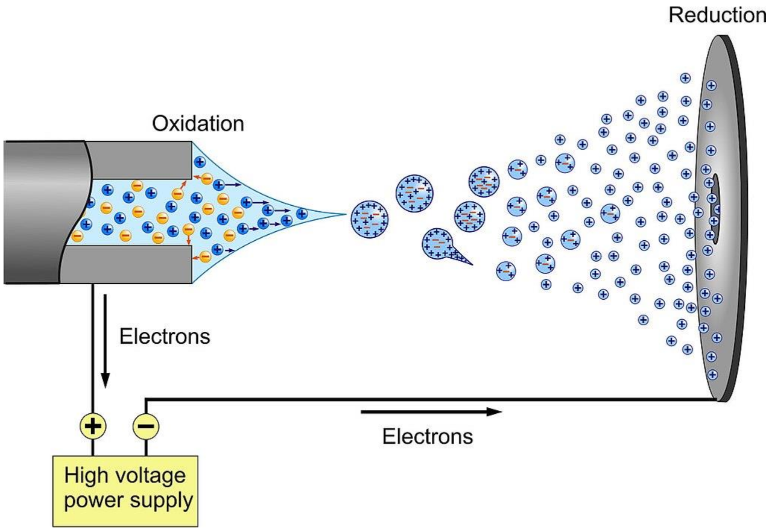


Fig. 2.10
Diagram of electrospray ionization in positive mode: under high voltage, the Taylor cone emits a jet of liquid drops. The solvent from the droplets progressively evaporates, leaving them more and more charged. When the charge exceeds the Rayleigh limit the droplet explosively dissociates, leaving a stream of charged (positive) ions.
en.wikipedia.org/wiki/Electrospray_ionization

This relation is shown in Fig. 2.8. Using this relation, one can find the "potential of zero charge" (PZC) for liquid electrodes (like the mercury electrode). At the PZC, the surface energy attains a maximum, so the drop comes closest to a spherical shape.

If the charge is sufficiently high, γ becomes negative and the surface area increases spontaneously. The result of this process is sometimes called a "Coulomb explosion." Applications of the Coulomb explosion include electro-spray ionization (ESI, Fig. 2.10) and electrospinning (Fig. 2.11). The process of electrowetting is related to this. In electrowetting, however, one stays below the threshold for the Coulomb explosion with the voltage and manipulates only surface energy and contact angle via the charge. Conductive droplets can be moved back and forth in an oil phase by a clever arrangement of electrodes and potential sequences on the surface ("droplet-based microfluidics"). One can also change the droplet radius and create switchable optical lenses.

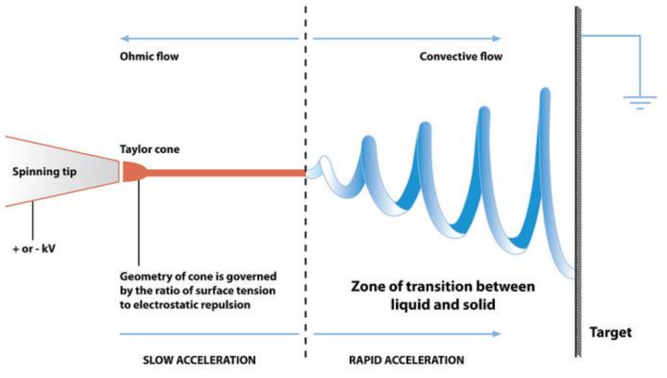


Fig. 2.11
To electrospinning
en.wikipedia.org/wiki/Electrospinning

Remember

- Charge migrates to the interface in conductive media. It lower the interfaces energy and can actually make it negative.
- When the surface energy becomes negative, a Coulomb explosion occurs. Applications are electro-spray-ionization, electrospinning, and electrowetting (the latter without Coulomb explosion).

2.5 Surface energies of solids

Solids also have a surface energy, which, however, cannot be measured easily because solids do not form droplets. Enlarging or diminishing the surface area of a solid always involves elastic work. The elastic work can be separated from the surface work, in principle, but the practicalities are cumbersome.

Solids usually have *high* surface energies. Tungsten, for example, has $\gamma = 4400$ mN/m. This is not surprising insofar, as tungsten – a “refractory metal” – has one of the highest melting points of all elements (3422°C).

A side remark: Most polymers have surface energies similar to those of liquids, although they appear to be solids. The low surface energy is a consequence of the chain topology. The cohesive forces *between chains* are similar to the cohesive forces in liquids.

Changes in the surface energy of solids upon adsorption can be tracked quite easily, making use of the bending thin bars ("cantilevers"). Sensors can be based on this bending. Fig. 2.12 shows a micromechanically fabricated beam.[21]

Solids can have heterogeneous surfaces. The surface energy then varies across the surface, as well. In consequence, the meniscus is not a straight, smooth line when the plate is pulled out of a water reservoir. The meniscus sticks to the defects. This is a simple test for cleanliness.

Remember

- Solids also have a surface energy. It is often high (> 1000 mN/m) because the internal cohesion is high.
- Measuring the surface energy of solids absolutely is difficult because elastic work is done when the interface area changes. *Changes of the surface energy* (e.g. caused by adsorption) can be inferred from the bending of thin beams.
- Solid-state interfaces often are laterally heterogeneous.

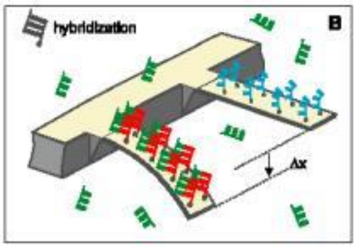


Fig. 2.12
Measuring surface energies of solids absolutely is difficult. The differential measurement of a *change of* γ can be done via beam bending at adsorption.
www.cs.unc.edu/~nanowork/tams.html

---

[21] Also "cantilever" (German: "Blattfeder"). Typical thicknesses of cantilevers are in the range of a few micrometers. Cantilevers are products of microsystems technology. They belong to the MEMS-devices (MEMS for “micro-electromechanical systems”) and have become cheap. In the academic field, bending is often read out with the “optical lever technique” (i.e. with a reflected laser beam). In engineering, this is considered too complicated. Mostly, piezoresistive elements are used on one side of the cantilever.

3 The capillary pressure

3.1 Motivation: The force exerted by capillary bridges

We begin the discussion of capillary pressure with an argument, which will prove to be incorrect, but which illustrates the usefulness of capillary pressure for calculations of all kinds. Consider the capillary bridge shown in Fig. 3.1. Such bridges give rise to a capillary force which pulls the two ends together. One might be tempted to calculate the capillary force as

Eq. 3.1.1

$$F_{cap} \overset{?}{=} \gamma S$$

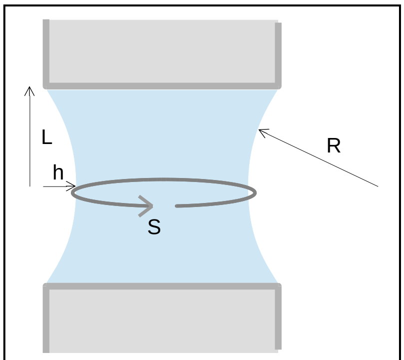


Fig. 3.1
A capillary bridge exerts an attractive force between the two solid surfaces. This force is *not* given as γS.

S is the circumference at the narrowest point (at the waist). Again, this formula is incorrect. It also contradicts experience. Insert values of $S = 1$ cm and $\gamma = 72$ mN/m. The force results as 0.7 mN, corresponding to a weight of $M = F_{cap}/g = F_{cap}/(9.81\ \text{m/s}^2) \approx 70$ mg. A capillary force of this magnitude would never be sufficient to support the weight of a microscope slide. The opposite is true. Two wet microscope slides adhering to each other strongly They can only be separated by a shearing motion. Pulling along the normal leads to nothing.

To correct the error in Eq. 3.1.1 we calculate the force as the derivative of work with respect to displacement:

Eq. 3.1.2

$$F_{cap} = \left(\frac{\mathrm{d}W_{cap}}{\mathrm{d}z}\right)_{Vliq}$$

At this point the old wisdom comes into play. When taking derivates in Physical Chemistry, it must always be specified which quantities are to be held constant. (Recall the difference between the specific heats at constant pressure and constant volume, c_p and c_V.) Eq. 3.1.1 would be correct if the circumference of the bridge at the narrowest point (at the "waist") remained constant during pulling. However, *the volume remains constant, instead,*[22] and the bridge becomes narrower at the waist (Fig. 3.2). The circumference decreases and the surface area A increases more strongly than what would be the case if the circumference was constant.

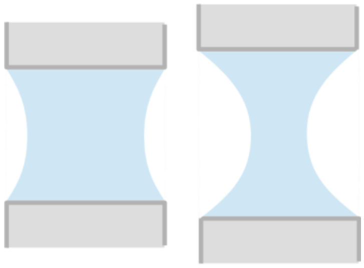


Fig. 3.2
When pulling on the solid bodies in Fig. 3.1 the liquid bridge constricts. This effect must be taken into account when calculating the surface work.

For the geometry from Fig. 3.2, the calculation of the surface area and its change when the distance between the two bodies is changed is tedious. It is easier to argue with the capillary pressure. The result from these considerations is

Eq. 3.1.3

$$F_{cap} = \gamma S - \Delta p\ A_{waist}$$

---

[22] Evaporation or condensation are left aside.

Δp is the capillary pressure (see below) and A_{waist} is cross sectional area of the bridge at the waist.

Remember

- The force exerted by a capillary bridge is *not* given as surface tension × circumference. The capillary pressure comes into play, as well.

3.2 *Capillary pressure, the Laplace equation*

Curved fluid-fluid interfaces result in a pressure difference between the two fluids. The difference is the *capillary pressure*, Δp. To calculate Δp, consider a fluid sphere (a drop or a bubble) and use that the free enthalpy, G, is at its minimum at equilibrium (Fig. 3.3). Let the radius of the sphere be such that G is minimal. Change the radius away from that value by a small amount dr. Two types of work are done, namely a pressure-volume work, given as $-\Delta p \mathrm{d}V_{drop}$, and the surface work, given as $\gamma \mathrm{d}A_{\mathrm{drop}}$ with A_{drop} the surface area of the droplet. Δp is defined as $p_{in} - p_{out}$ and V_{drop} is the volume of the drop, not to be confused with the volume of the entire system. Since both V_{drop} and A_{drop} depend on r, the chain rule needs to be applied:

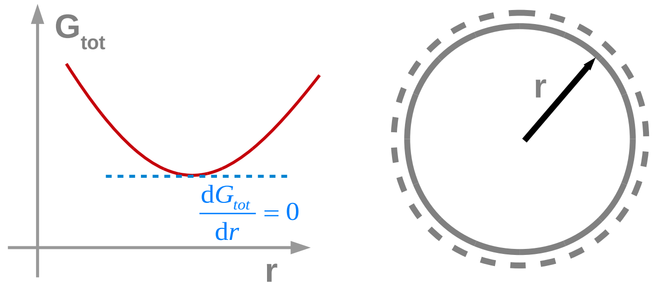


Fig. 3.3
To calculate the capillary pressure, consider a sphere of a fluid in thermodynamic equilibrium. If the radius of this sphere is changed by a small amount dr, the work done in the process must be zero, in total.

Eq. 3.2.1

$$\frac{\mathrm{d}G_{tot}}{\mathrm{d}r} = -\Delta p \frac{\mathrm{d}V_{drop}}{\mathrm{d}r} + \gamma \frac{\mathrm{d}A_{drop}}{\mathrm{d}r} = 0$$

We insert the volume and the surface area of a sphere:

Eq. 3.2.2

$$-\Delta p \frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{4}{3}\pi r^3\right) + \gamma \frac{\mathrm{d}}{\mathrm{d}r}\left(4\pi r^2\right) = 0$$

Differentiation results in

Eq. 3.2.3

$$-\Delta p\ 4\pi r^2 + 8\pi\gamma r = 0$$

and

Eq. 3.2.4

$$\Delta p = \frac{2\gamma}{r}$$

This is the Laplace equation. Δp is also called „Laplace pressure“[23].

The calculation leading to Eq. 3.2.4 assumes spherical shape. For a long cylinder ($L >> r_{\mathrm{cyl}}$), one has $V_{\mathrm{cyl}} = \pi r_{\mathrm{cyl}}^2 L$ and $A_{cyl} = 2\pi r_{\mathrm{cyl}} L$. The capillary pressure then is $\Delta p = \gamma / r_{\mathrm{cyl}}$.

Eq. 3.2.1 can be rearranged as

Eq. 3.2.5

$$\Delta p = \gamma \frac{\mathrm{d}A_{drop}}{\mathrm{d}V_{drop}}$$

For arbitrarily shaped interfaces, this leads to

Eq. 3.2.6

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

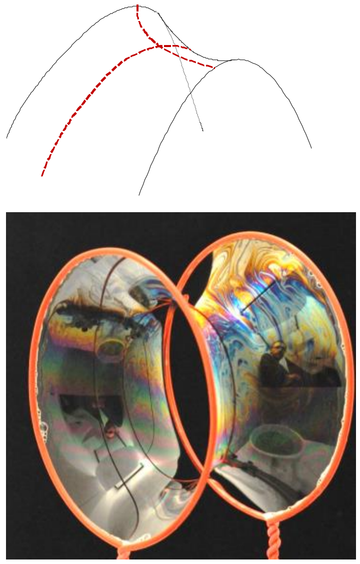


Fig. 3.4
A minimal surface
www.soapbubble.dk/english/science/the-geometry-of-soap-films-and-soap-bubbles/

$1/R_1$ and $1/R_2$ are the "principal curvatures". For the explanation of the principal curvatures we borrow from differential geometry.[24] On any point of a curved surface, one can find lines which have no curvature in the plane. They only are curved perpendicularly to the surface. These lines are called “geodesics”. The geodesics are at the same time the shortest connections between two points. At the north pole, the geodesics are the longitudes. Among the geodesics, there is always one with a maximum out-of-plane curvature and one with minimum out-of-plane curvature. Of course, the out-of-plane curvatures may all be the same. That is so at the north pole. It is not so at the equator because the earth has oblate shape. Among the geodesics are the equator (with the smallest curvature, that is, the largest radius of curvature) and the longitude (with the largest curvature). The two geodesics with maximum and minimum curvature are perpendicular to each other. $1/R_1$ and $1/R_2$ are the maximum and minimum curvature. Back to the capillary pressure: If the curved surface is not a segment of a sphere, the Laplace equation contains the term $(1/R_1 + 1/R_2)$ instead of the term $2/r$. The parameter $(1/R_1 + 1/R_2)$ is sometimes called “mean curvature”.[25]

The surface of a body (a drop, a soap bubble, ...), which has its surface area minimized under a given boundary condition (attachment to a solid surface, fixed volume....), has the same mean curvature everywhere ($1/R_1 + 1/R_2 = \mathrm{const.}$). That can be proven mathematically. From the viewpoint of Physical Chemistry, this simply says that the capillary pressure is the same everywhere.

---

[23] An opinion on terminology: If there is a choice for naming something, naming it after the discoverer (here: Laplace) is the second best choice. Speaking terms (here: capillary pressure) are better terms.

[24] Differential geometry is the geometry of curved surfaces or – more generally – the geometry in curved spaces. Einstein moved differential geometry from a niche of mathematics to the center of how we view the world by claiming that space-time was curved. In general relativity, inertial systems move on geodesics in four-dimensional space-time.

[25] “Mean curvature” occasionally also is defined as $(1/R_1 + 1/R_2)/2$.

A word of caution: If there is a transport equilibrium between liquid and vapor in porous media, the capillary pressure mostly depends on the relative humidity rather than the geometry. This is so because capillary condensation (section 4) governs the radius of curvature.

Eq. 3.2.6 has the interesting consequence that the capillary pressure can vanish even if the surface is not planar. For these particular surfaces, the two principal curvatures point in opposite directions. Fig. 3.5 shows an example. One may create such *minimal surfaces* by dipping a loop of a wettable wire into a vessel filled with water and soap, followed by pulling the loop out in a way that leaves a soap lamella attached to the wire. The capillary pressure is zero because the pressure is the same on both sides of the lamella. Surfaces with vanishing mean curvature are the smallest possible surfaces bounded by the edge (hence the name "minimal surface"). That the minimal surfaces have vanishing mean curvature everywhere ($1/R_1 + 1/R_2 \equiv 0$) is proven in differential geometry with some effort. If the Laplace equation is expressed in the form of Eq. 3.2.6, this result is plausible without mathematics.

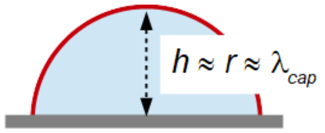


Fig. 3.6
If the characteristic length of a problem (e.g. the radius of a droplet) is equal to the capillary length, gravitational forces and capillary forces roughly balance each other. On smaller length scales capillarity predominates, on larger ones gravitational forces.

We estimate the magnitude of capillary pressure for nanodroplets and nanobubbles. Assume $\gamma = 50$ mN/m and $r = 10$ nm. Inserting these values into Eq. 3.2.4 leads to $\Delta p \approx 10^7$ Pa = 100 bar. The capillary pressure inside nanosized drops is substantial. It is so substantial that nanobubbles are not easily stabilized, mechanically. If no special measures are taken (a metallic matrix or surface tension close to zero) they collapse. Nanofoams are challenging from a materials engineering perspective.[26] However, such large pressures only occur when the droplets or bubbles are indeed of nanoscopic dimension. The capillary pressure of a drop of water on a car's windshield (let the radius be 1 mm), is only 10^{-3} bar. There is a *scale dependence*. Capillary pressure is only significant on small scales. Another example: an air bubble in a vein ($r \approx 10$ µm) is equivalent to a plug (which is dangerous).

If the capillary pressure competes with a pressure of some other origin, the size dependence of the capillary pressure and the competing pressure results in a characteristic length, which is the *capillary length*. Consider a droplet sitting on a plate (Fig. 3.6). It might also be a pending drop (Fig. 3.7) or it might sit on a vertical surface (like a raindrop on a window pane). At what drop size will gravity noticeably affect this drop's shape? Experience tells us that this critical size is about 1 mm. Droplets larger than that slide down the window pane. The value of this critical size can be understood by demanding that the capillary pressure and the gravitational pressure are of similar magnitude:

Eq. 3.2.7

$$\frac{2\gamma}{r} \approx \rho g h$$

$g = 9.81$ m/s^2 is the acceleration constant due to gravity and h is the size of the drop. Assume $h \approx r$ and call this length λ_c (the capillary length). Solving for λ_c results in

Eq. 3.2.8

$$\lambda_c \approx \sqrt{\frac{\gamma}{\rho g}}$$

---

[26] There are also foams in which the bubbles are separated only by lamellae.

A factor of 2 has been omitted. For water, $\lambda_c \approx 2$ mm. (Sometimes the factor 2 is not omitted. The capillary length then has a slightly different value).

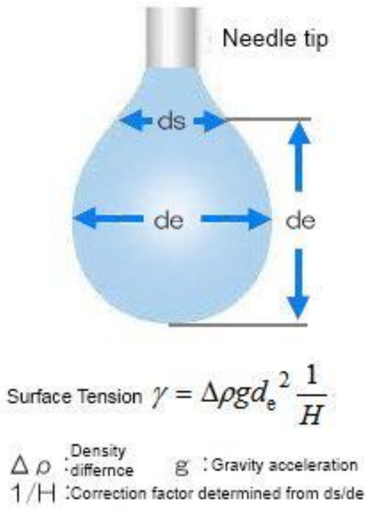


Fig. 3.7
For measuring surface tension using the pendant drop method. *www.face-kyowa.co.jp/english/en_science/en_theory/en_what_Surface_tension/*

If the drop size is fixed, the ratio of gravitational energy to the surface energy sometimes is called "Bond number".[27] Drops which are larger than the capillary length have a Bond number larger than 1.

The capillary length underlies a rather accurate method for determining the surface energy. This is the *pendant drop method* (Fig. 3.7). The contour of a pendant drop is captured by a camera and analyzed by a computer. If the drop size is comparable to the capillary length, the surface tension can be derived from the contour. The lower the drop hangs, the smaller the surface tension.

Characteristic lengths similar to the capillary length (and often also called "capillary length") are also used in other contexts. In these situations, the capillary pressure always is in competition with another pressure. An example is the splitting up of emulsion droplets in a shear flow.[28] Here the capillarity competes with the stress exerted by viscous medium. We leave details away. The smaller the droplet, the more difficult it is to break it up in the flow field. A practical lower limit is in the range of $r \approx 50$ µm.[29]

Sometimes such competition is also quantified by a *capillary number* (*Ca*). Consider again the fragmentation of droplets in the shear field. The capillary number is defined as

Eq. 3.2.9

$$Ca = \frac{\eta v}{\gamma} = \frac{\eta v}{r}\frac{r}{\gamma}$$

η is the viscosity[30] and v is a characteristic velocity. The parameter $\eta v/r$ is a viscous stress and γ/r is the capillary pressure (up to a factor of 2). Eq. 3.2.9 does not explicitly refer to the drop size. The size dependence of *Ca* arises because the "characteristic velocity", v, for this case is the product of shear rate and droplet size. If the capillary number is large (that is, if the drop is large enough), the drop can be fragmented by the shear field.

Fig. 3.8
The pressure inside a liquid (dark blue) can be negative if the liquid is contained in a pore and if the liquid-air interface is concave.

There are other definitions of the capillary number, adapted to other situations. The surface tension is always in the denominator. If $Ca \gg 1$, capillarity can be overcome by the competing force.

Capillary pressure can cause the local hydrostatic pressure to be negative in condensed matter. (This is impossible in a vacuum or in gases.) Consider the situation shown in Fig. 3.8. A notch in the

---

[27] There are other, slightly different definitions of bond number.
[28] An emulsion is a mixture of two liquids that are not "molecularly dispersed".
[29] In this area, one must always keep a careful eye on whether the radius or the diameter is used to indicate the drop size.
[30] For accurate treatment, one must take into account that the droplet and the surrounding fluid generally have different viscosities.

surface of a solid contains a liquid. The solid is wetted by the liquid. If the (negative) capillary pressure is larger than the ambient pressure (1 bar), the absolute pressure is negative inside the liquid. This is a consequence of the cohesive forces and the rigid wall. Negative capillary pressure is common inside wet porous media. This includes drying latex films. The negative pressure deforms the latex particles (Sect. 11.2).

Negative capillary pressure also explains the "Schröder paradox". Hydrogels swell more in liquid water than they swell in water-saturated air. When dipping a cylinder of a hydrogel into water such that is immersed half way, the cylinder swells more below the water surface than above, even though the chemical potential of H_2O is the same on both sides. The air-water interface at the outside of the cylinder in air is not flat on the nano-scale. Its nano-scale curvature contracts the gel.

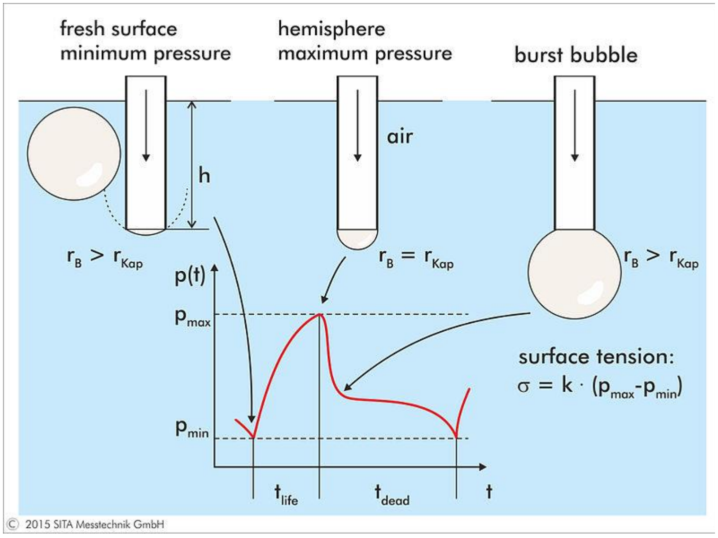


Fig. 3.9
The bubble pressure method
commons.wikimedia.org/wiki/File:Bubble_pressure_method_-_surface_tension.jpg

Remember

- For spheres, the capillary pressure is given as $\Delta p = 2\gamma/r$.
- The capillary pressure is large when the radius of curvature is small. Nanobubbles and nanodroplets are under a pressure of a few bars.

3.3 Consequences of the capillary pressure

The consequences of capillary pressure are numerous. We discuss a few examples:

- The surface tension of liquids can be determined by the bubble pressure method (Fig. 3.9). The end face of the capillary must be designed such that the three-phase boundary line adheres to the outer edge. If one starts from the situation in Fig. 3.9 on the left, and air is forced through the capillary, the radius of curvature initially *decreases.* This situation is stable because the pressure increases with increasing bubble volume. This changes when the radius of curvature becomes larger than the radius of the face. Then the pressure decreases with increasing bubble volume and *instability* occurs . The bubble ruptures. The maximum pressure is given as $p_{max} = 2\gamma/r_S$ with r_S the radius of the opening of the tube. From p_{max}, one calculates the surface tension.

- Capillary bridges in conjunction with negative capillary pressure lead to *capillary adhesion*. Capillary adhesion is well known from sand piles. Also, some animals secrete a fluid when walking on smooth surfaces. The fluid lets the feet be sticky.
  At this point, we remind ourselves of the force exerted by a capillary bridge (Sect. 3.1). The capillary force is *not* given as γS with S the circumference. The capillary force is given as $-\Delta p A + \gamma S$. A and S are to be evaluated at the waist. (There are extensions of this equation for geometries other than the bridge from Fig. 3.1.) Note that $\Delta p < 0$ for capillary bridges (most of the time). For macroscopic situations, the term $-\Delta p A$ governs the force. Capillary adhesion is also very relevant at the microscopic level and at that level, the term γS must not be ignored.

- Attractive capillary forces create a tensile stress when porous media dry out. This can lead to cracking. Structures with high porosity often collapse during drying. (That can be avoided with supercritical fluids.)
- *Capillary rise* occurs in hydrophilic tubes (Fig. 3.11). Equating capillary pressure and the pressure caused by gravity ($2\gamma/r = \rho g h$) leads to

Eq. 3.3.1

$$h = \frac{2\gamma}{r\,\rho\,g}$$

When equating the radius of the tube, *r*, with the radius of curvature, it was assumed that the contact angle is zero. If the contact angle θ is finite, the radius *r* must be replaced by $r/\cos\theta$ (see the right-hand side in Fig. 3.10).

Note that the height of the capillary rise does not depend on the thickness of the wall. This was first observed in 1808 by Clairaut. Clairaut concluded that the underlying forces are of short range. Using the parameters of water ($\rho = 10^3$ kg/m^3, $g = 9.81$ m/s^2, $\gamma = 72\times10^{-3}$ J/m^2), a contact angle of 0°, and a radius of the capillary of 1 µm, a height of $h = 7$ m is obtained. This value is similar to the height of trees. Trees use capillarity to pump fluid from the roots to the leaves. This argument is not quite as simple as it sounds. The capillary radius only needs to be small at the leaves. Elsewhere, the capillaries may be wider and they are in fact wider because wider capillaries have less resistance to flow.

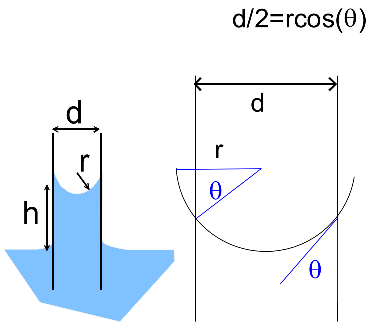


Fig. 3.10
When the contact angle in a capillary is small, the water rises in the capillary.

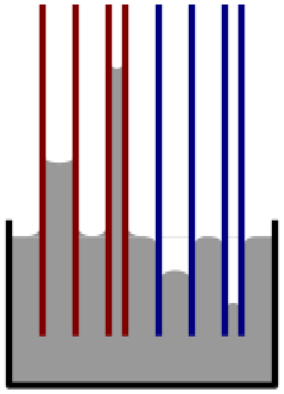


Fig. 3.11
The rise height depends on the radius of the capillary. Hydrophobic surfaces lead to capillary depression for water.

- Capillary pressure may drive the transport of liquids in porous media. This is important in printing on paper ("imbibition"), in candles ("wicking"), and in agriculture (take-up of rain by the soil). The same mechanism operates when a piece of sugar is dipped into a coffee mug. Capillary forces also drive transport in the "lateral flow device" (an "assay", an example is the Covid19 test strip).
The simplest equation to describe the imbibition is the Lucas-Washburn equation

Eq. 3.3.2

$$z(t) = \sqrt{\frac{\gamma r \cos\theta}{2\eta}\,t}$$

To prove this, equate the capillary pressure to the pressure caused by the hydrodynamic drag. These pressures are:

Eq. 3.3.3

$$\text{I: } \Delta p_{stream} = \frac{8}{r^2}\eta\, z(t)\frac{dz}{dt} \qquad \text{II: } \Delta p_{cap} = \frac{2\gamma\cos\theta}{r}$$

Equation I is the Hagen-Poisseuille law. *z* is the length to which the capillary is filled. dz/dt is the flow velocity. Use $\Delta p_{stream} = \Delta p_{cap}$, separate the variables, and integrate:

Eq. 3.3.4

$$\int_0^z z'(t')\,\mathrm{d}z' = \int_0^t \frac{\gamma r \cos\theta}{4\eta}\,\mathrm{d}t'$$

The tildes were introduced so that the integration variables are named differently from the integration limits. One finds

Eq. 3.3.5

$$\frac{1}{2}z^2 = \frac{\gamma r \cos\theta}{4\eta}t$$

Solving for z proves the Lucas-Washburn equation (Eq. 3.3.2). Eq. 3.3.2 describes transport in a capillary. For systems of pores there is a set of complications and a set of extensions dealing with the complications.

A side remark: z scales like $t^{1/2}$, similar to the root-mean-square displacement in diffusion. However, this is *not* diffusion in the sense of Brownian motion.

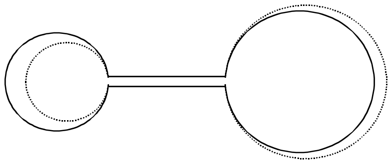


Fig. 3.12
When two soap bubbles of unequal size are in contact with each other, the material flows to the larger bubble. This can be explained by the capillary pressure (Eq. 5.2.6). It can also be explained by the decrease in total surface area.

A subtlety: According to the Lucas-Washburn equation, the large pores fill first. Later, the liquid is redistributed towards the small pores, because the negative capillary pressure is larger there.

Remember

- The pendant drop method takes advantage of a competition between capillary pressure and gravity to determine surface tension.
- Capillary pressure can lead to capillary adhesion.
- Capillary pressure leads to capillary rise. This drives the liquid transport to the tops of trees.
- Capillary pressure may drive fluid transport in porous media.

3.4 Capillary instabilities

Consider a tube with a valve, with soap bubbles attached to both sides (Fig. 3.12). Let the radius of the bubble on the left (r_L) be smaller than the radius of the bubble on the right (r_R). After the valve is opened, the small bubble will become smaller and the large bubble will become larger. The situation is *unstable*.

At this point, capillary pressure is often used as an argument. Because the capillary pressure is larger on the left than on the right, the left bubble shrinks. The instability can also be understood from the change of the total surface area. For the bubbles shown in Fig. 3.12, the sum of the surface areas of both spheres decreases if the sum of the volumes remains constant and the smaller sphere transfers its volume to the larger sphere. Take advantage of the constant volume:

Eq. 3.4.1

$$\frac{4\pi}{3}\left(r_A^3 + r_B^3\right) = const \quad \Rightarrow r_A^2\mathrm{d}r_A = -r_B^2\mathrm{d}r_B$$

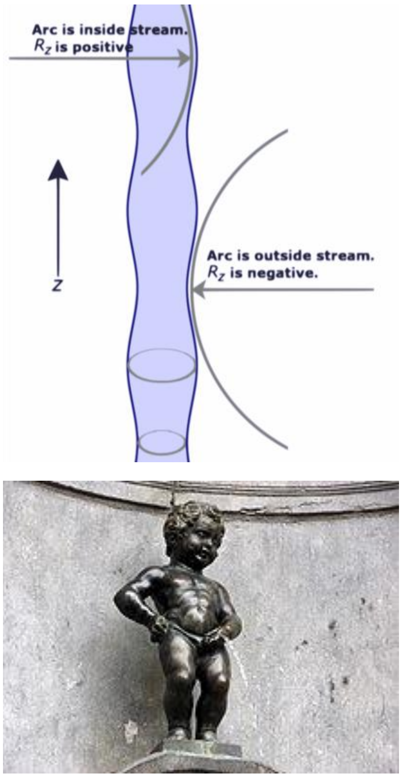


Fig. 3.13
The Rayleigh instability
en.wikipedia.org/wiki/Surface_tension
en.wikipedia.org/wiki/Urination

Calculate the change of the surface area as

Eq. 3.4.2

$$\mathrm{d}A = 4\pi\left(\mathrm{d}\left(r_A^2\right) + \mathrm{d}\left(r_A^2\right)\right) = 8\pi\left(r_A \mathrm{d}r_A + r_B \mathrm{d}r_B\right)$$

$$= 8\pi r_A \mathrm{d}r_A \left(1 - \frac{r_B^3}{r_A^3}\mathrm{d}r_B\right) < 0 \quad \text{if } r_B > r_A$$

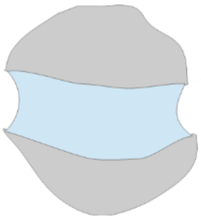


Fig. 3.14
A pore filled with a fluid. If both menisci are concave, the situation is stable.

A similar situation is known for liquid cylinders (Fig. 3.13). If – for some reason (a thermal fluctuation?) – the cylinder is slighter thinner somewhere, the capillary pressure is increased there and the surface tension further contracts this part of the cylinder until the cylinder eventually splits into drops. This is the *Rayleigh instability*. An implicit assumption in this argument is that the curvature of the surface along the axis is less than across the axis. Thus, the unstable perturbations must have a certain minimum length. On the other hand, very long wavelength perturbations require material transport over long distances to turn the jet into a chain of droplets. From these two competing requirements, one calculates a characteristic droplet size. The details are even more complicated. High-speed videos show that a droplet in the process of pinching off breaks into more smaller droplets. The dynamics is complicated but reasonably well understood. That there is more than a single characteristic drop size can also be seen in the spider webs, on which dew has formed drops overnight. One sees drops of different sizes. These drops are also a consequence of the fact that a uniformly thick liquid film on a thin fiber is unstable.

Note: Two menisci in hydrostatic equilibrium are stable if the concave (Fig. 3.14).

<u>Remember</u>

- When two soap bubbles of different sizes are in contact with each other, the larger one grows at the expense of the smaller one.
- A related instability (the Rayleigh instability) causes liquid cylinders to decay into droplets.
- These instabilities can be understood based on either capillary pressure or the lowered total surface area.

3.5 Ostwald ripening

Another important instability is Ostwald ripening. Replace the flow through the tube in Fig. 3.12 by a transport through the vapor phase.[31] There then is an evaporation-condensation equilibrium at the surfaces of the droplets. (The local vapor pressure is larger over the small droplets than over the large ones, Sect. 4.1.) Diffusion through the continuous phase causes the large droplets to grow at the expense of the small droplets.

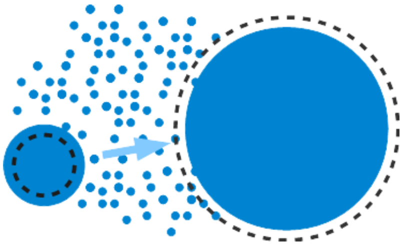


Fig. 3.15
In Ostwald ripening, large drops grow at the expense of small drops.

Ostwald ripening is a widespread phenomenon. It turns fog drops into rain drops. It also occurs in emulsions. In this case, there is some residual solubility of the oil in the water phase. The dissolved molecules migrate from the small drops to the large drops. Emulsions must be stabilized against a) sedimentation, b) coalescence and c) Ostwald ripening. Without

---

[31] The same argument applies to oil droplets in water, if there is some residual solubility of the oil in the water. It also applies to air bubbles in a liquid. We formulate the facts for droplets in steam; the transfer to the other cases is simple.

stabilization, the oil phase will separate from the water phase (as is known from milk).

For emulsions, Ostwald ripening can be avoided by adding a substance to the oil phase, which is poorly soluble in the water phase. When the droplets shrink, the concentration of this substance increases. It builds up an osmotic pressure and thus brings Ostwald ripening to a halt. In the context of *miniemulsions*, this substance carries the name " hydrophobe". (These are oil-in-water emulsions.) A typical hydrophobe is hexadecane. "Miniemulsions" are not characterized by a certain size range (e.g., to be distinguished from microemulsions[32]). Miniemulsions[33] are characterized by this particular mechanism of stabilization.

Ostwald ripening also occurs in polycrystalline solids. In these, large crystallites grow at the expense of small crystallites. Sufficiently high temperature is required (meaning, a temperature close to the melting point). Controlling Ostwald ripening is part of the “thermomechanical treatment" of steel. The material is heated, forged, and rolled at the same time. Crystallites become larger and the arrangement of grain boundaries is manipulated (by forging or rolling). Remember that capillary instabilities are ultimately driven by the surface-to-volume ratio (see the discussion before Eq. 3.4.2). Polycrystalline materials can lower the total energy of their inner surfaces by Ostwald ripening.

Remember

- In Ostwald ripening, large droplets (bubbles, particles) grow at the expense of small ones.
- Ostwald ripening is similar to a capillary instability.

---

[32] Microemulsions contain surfactant mixtures that cause spontaneous curvature. Microemulsions are thermodynamically stable. They form spontaneously. Other emulsions are metastable. Microemulsions typically have droplet radii in the range of a few nm. They require large amounts of surfactant, which is a disadvantage.

[33] The term "miniemulsion" is mainly used in the context of miniemulsion polymerization. Stabilization of emulsions against Ostwald ripening by addition of long-chain non-polar molecules is also used in the food sector, in crop protection, and in cosmetics.

4 Vapor pressure above curved surfaces, nucleation

4.1 *The Kelvin equation*

Consider the two liquids in Fig. 4.1, which differ in curvature of the surface. The saturation vapor pressure is higher above the convexly curved surface, as can be understood by equating chemical potentials in the liquid phase and the vapor phase:

Eq. 4.1.1

$$\text{Ia: } \mu_{liq} = \mu_{vap} \qquad \text{IIa: } \mu_{liq}' = \mu_{vap}'$$

$$\text{Ib: } \mu_{liq} = \mu_{vap}^{\varnothing} + RT \ln \frac{p_{sat}}{p^{\varnothing}} \qquad \text{IIb: } \mu_{liq} + \bar{V}_{liq}\Delta p_{hyd} = \mu_{vap}^{\varnothing} + RT \ln \frac{p_{sat}'}{p^{\varnothing}}$$

Primed parameters refer to the situation on the left in Fig. 4.1. In equation Ib, ideal behavior was assumed for the chemical potential in the gas phase. In equation IIb, it was used that the liquid is pure. Take the difference IIb – Ib:

Fig. 4.1
The saturation vapor pressure is increasd above convexely curved surfaces.

Eq. 4.1.2

$$\bar{V}_{liq}\Delta p_{hyd} = RT\left(\ln \frac{p_{sat}'}{p^{\varnothing}} - \ln \frac{p_{sat}}{p^{\varnothing}} \right)$$

which leads to

Eq. 4.1.3

$$\ln \frac{p'_{sat}}{p_{sat}} = \frac{\bar{V}_{liq}\Delta p_{hyd}}{RT}$$

The hydrostatic pressure thus causes the vapor pressure to increase. Figuratively speaking, the increased hydrostatic pressure drives the molecules out of the liquid into the vapor. One uses the capillary pressure for the hydrostatic pressure and obtains

Eq. 4.1.4

$$\ln \frac{p'_{sat}}{p_{sat}} = \frac{\bar{V}_{liq} 2\gamma}{RTr}$$

p_{sat} and p_{sat}' are the equilibrium vapor pressure over a planar surface and over a curved surface, respectively. The subscript "*sat*" stands for saturation. Eq. 4.1.4 is the *Kelvin equation*. Increased capillary pressure causes increased vapor pressure above that surface.

4.2 *Homogeneous nucleation*

The Kelvin equation is closely related to *homogeneous nucleation*. When the saturation vapor pressure is exceeded in a vapor phase (as often occurs during cooling, including summer evenings), this often leads to a *metastable* situation because nanoscopic droplets decay shortly after having been formed. The equilibrium vapor pressure above a nanodroplet is much higher than the equilibrium vapor pressure above a planar surface. Only when a certain critical droplet size is exceeded, do the droplets grow. One can calculate the critical drop size from the Kelvin equation by rearrangement:

Eq. 4.2.1

$$r_{crit} = \frac{\bar{V}_{liq}\,2\gamma}{RT\ln\left(p'_{sat}/p_{sat}\right)} = \frac{\bar{V}_{liq}\,2\gamma}{RT\ln\left(rH\right)}$$

rH stands here for relative humidity. (The last transformation only makes sense for water.) The larger the supersaturation, the smaller the critical radius.

In the following, we discuss nucleation not on the basis of the Kelvin equation, but on the basis of volume energy and surface energy. This description is equivalent to the reasoning with the Kelvin equation (which we do not prove). The equivalence echoes the treatment of capillary instabilities based on either of capillary pressure or the decrease in total surface area.

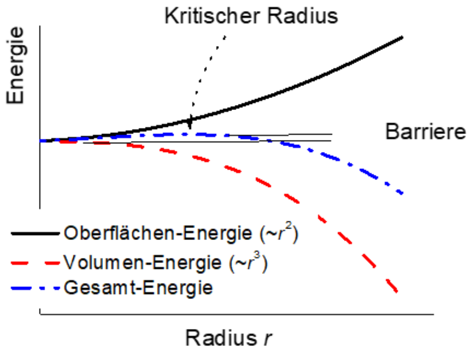


Fig. 4.2
Homogeneous nucleation can be understood from plotting the sum of the total surface energy ($\propto r^2$) and the total volume energy ($\propto r^3$) versus r.

Fig. 4.2 has the radius of a droplet on the x-axis. The total energy of the particle has two contributions: a surface energy and a volume energy. The surface energy is quadratic in r (is $4\pi r^2$ for spheres). The coefficient is the surface energy γ. The volume energy is cubic in r (is $4\pi/3\ r^3$ for spheres).[34] The coefficient is the difference of the chemical potentials in the outer space and in the particle, normalized to molar volume $\Delta\mu/\bar{V}$. $\Delta\mu$ increases with increasing supersaturation. When subtracting a cubic function from a parabola, one finds a maximum at small *r*, because the quadratic term in the limit of $r \to 0$ is always larger than the cubic term (even if both terms approach zero, individually). In consequence, there always is a nucleation barrier. The critical radius (the radius at the barrier) is the same as the critical radius as derived from Eq. 4.2.1. The proof is a bit tedious. The derivation via the energies also allows to estimate the height of the nucleation barrier.

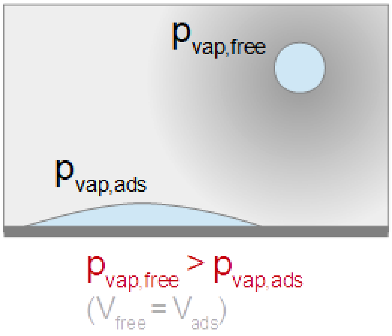


Fig. 4.3
Heterogeneous nucleation of liquid droplets occurs at surfaces which are wetted by the liquid because the radius of curvature of the liquid-vapor interface becomes larger (compared to spherical drop with the same volume).

An alternative to homogeneous nucleation is heterogeneous nucleation, i.e. nucleation at surfaces. There is an important difference between liquids and crystals. For liquids, the lowered nucleation barrier of heterogeneous nucleation is the consequence of the larger radius of curvature (assuming the same volume and a hydrophilic surface, Fig. 4.3). In crystals, *epitaxial growth* is a second requirement. Heterogeneous nucleation of crystals requires a surface of the same crystal or at least a surface with similar structure (similar spacing between the lattice plains, in particular).

4.3 Capillary condensation

If fluid-gas interface is curved towards the gas phase (if it is "concave", Fig. 3.8), the radius in the Kelvin equation (Eq. 5.2.4) is formally negative. Liquid condenses onto such surfaces even at *undersaturation*. The process is called *capillary condensation*. Because of capillary condensation,

---

[34] This argument would apply in a similar same way if r were the size of a cube. In this case, the surface area would be proportional to $6r^2$ and the volume would be proportional to r^3. The surface-to-volume ratio scales as the inverse size, regardless of shape. (An exception are the "fractal" objects.)

hydrophilic porous media and hydrophilic rough surfaces often are wet. Porous media take up water, thereby turning into drying agents. A well-known example is silica gel.

At contacts, capillary condensation leads to *capillary adhesion* (Fig. 4.4). Examples are:

- Sand piles are stabilized by capillary bridges. (Sand castles on the beach are also glued together by dry salt. That is different from capillary condensation).
- Capillary forces play a role in atomic force microscopy. Dry air is needed to avoid these.
- Capillary condensation is one reason for why the static friction coefficient (sec. 9.2) increases with increasing rest time. The phenomenon was already reported by Leonardo da Vinci. It was long thought that the main cause of this increase was plastic deformation of asperities[35] in the contact zone ("asperity creep"). “Capillary aging” also plays a role.

As shown in Fig. 4.4, capillary condensation is often hysteretic because a capillary bridge, once formed, does not break again. Upon drying out, it pulls the two surfaces together and the surfaces remain stuck to each other. This interplay of capillary forces and other forces is of much importance in soil mechanics.

A side remark on capillary forces. At the beginning of section 3.1 it was reported that the capillary force can be calculated from the virtual work and, further, that the virtual work must be calculated at constant fluid volume (as opposed to constant circumference at the waist). At the nanoscale transport in the vapor phase is fast. For all liquids with a substantial vapor pressure (for all liquids except ionic liquids and silicone oils), the boundary condition of the virtual work on nanosized droplets is *constant chemical potential of the liquid* rather than constant volume. Constant chemical potential implies constant mean curvature. One might think that this change of the boundary condition would have an influence on the capillary force. In calorimetry, c_p is different from c_V. Constant volume or constant pressure do make a difference for the heat capacity. However, the capillary force does not depend on whether the volume or the chemical potential are constant. This is so because an energy is linked to evaporation or condensation when pulling onto the capillary bridge at constant chemical potential. This energy exactly compensates the difference between the two surface works, calculated for constant volume or constant chemical potential.

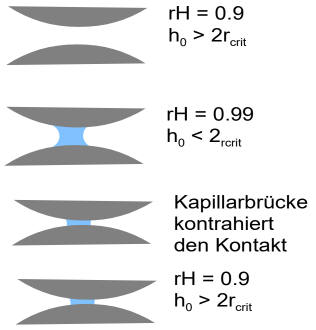


Fig. 4.4
Capillary adhesion is hysteretic when a bridge, once formed, does not reopen.

Remember

- Above convexly curved surfaces, the vapor pressure is higher than above planar surfaces.
- The vapor pressure follows the Kelvin equation: $\ln(p_{sat}'/p_{sat}) = (2\gamma/r)/(RT/V_{mol})$.
- There is a critical minimum droplet size for droplet growth in supersaturated vapor. Once this radius is reached, homogeneous nucleation sets in.
- For droplets on wettable surfaces, the radius of curvature can be larger than in the bulk (assuming identical volumes), which leads to heterogeneous nucleation.
- The mechanisms of homogeneous and heterogeneous nucleation can also be understood from the sum of surface and bulk energies plotted versus radius.
- Homogeneous and heterogeneous nucleation also exists for crystals, although these do not have spherical shape. Nucleation can also be explained by the surface-to-volume ratio. Heterogeneous nucleation of crystals requires epitaxial growth of the crystal phase on the surface.

---

[35] “Asperities” are protrusions of a rough surface.

- Above concave surfaces, the saturation vapor pressure is lowered. This leads to capillary condensation even in undersaturated vapor. Capillary condensation is of much importance because it leads to capillary adhesion.

5 Adsorption

5.1 *General remarks*

Adsorption of molecules at interfaces is important first because it is ubiquitous. Whenever adsorption lowers the surface energy, it takes place. (It is the *free* energy of the interface. Entropy can also drive adsorption, although these effects are usually rather small.[36]) The relationship between energy of adsorption and adsorbed amount is the content of the Gibbs adsorption isotherm.

Important applications of adsorption are

- Surface modification
- Adhesion
- Adhesion of coatings to their substrate
- Corrosion protection
- Compatibilization
- Flotation
- Adsorption on activated carbon: purification of air
- Chromatography and inverse chromatography
- Surface-based sensing

The prevention of adsorption is as important in technology as adsorption itself. The *cleaning of* surfaces cannot be overestimated in its importance. Because of ubiquitous adsorption, it is virtually impossible to keep metal surfaces clean.[37] This includes electrodes. The key advantage of the mercury drop electrode in electrochemistry is that it is constantly regenerating.[38] Silicones are good biomaterials in part because they have low surface energy.[39]

In adsorption, one must always distinguish between the adsorption *kinetics* and the adsorption *isotherm.* The former is the adsorbed amount as a function of time, the latter is the adsorbed amount at equilibrium as a function of the concentration (more precisely, of the activity) in the adjacent bulk phase. A practical problem: The adsorption equilibrium requires time. That is particularly important for strong adsorption to solid surfaces, because the adsorbate molecules often reorganize slowly. When an adsorption equilibrium is out of reach, one sometimes resorts kinetic measurements and evaluates the rate constants of adsorption and desorption, k_{on} and k_{off}. One may also have a model for adsorption kinetics, from which one can derive the adsorbed amount at equilibrium, based on this this kinetic model. Fig. 5.2 shows an example.

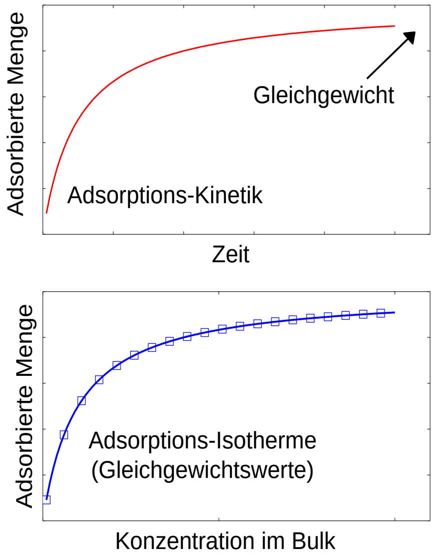


Fig. 5.1
An adsorption isotherm shows the equilibrium values of the adsorbed amount versus the concentration in the bulk. These values are the final results of many measurements of adsorption kinetics at different bulk concentrations. Depending on how quickly equilibrium is reached, one can also ramp the bulk concentration as a slow ramp and hope to be close to adsorption equilibrium at all times.

---

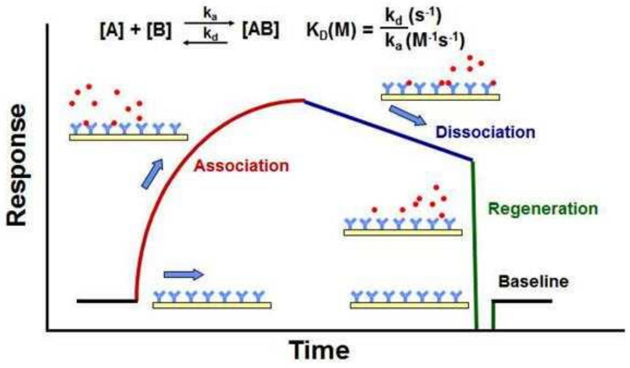
[36] When entropy dominates, large molecules (including proteins or polymers, for example) displace small molecules at an interface because the many small molecules lose less entropy during adsorption than the few large molecules. Small molecules in this context are often solvent molecules

[37] Clean metal surfaces have a high surface energy (but they are - as said - almost never clean).

[38] The surface of a jet of a liquid directly behind the outlet nozzle is also relatively clean.

[39] "Low-fouling" surfaces are also efficient if they bind water well. In liquids, adsorption is always *competitive adsorption*.

The measurement of rate constants is a problem when the adsorption kinetics is transport-limited. Remember the unstirred layer of the thickness of the Nernst length (a few 100 µm), inside which all transport is diffusive. Diffusive transport implies a time constant $\tau_{diff} = L^2/D$ with L the thickness of the Nernst layer (a few 100 µm) and D the diffusivity ($D \approx 10^{-5}$ cm²/s = 10^{-9} m²/s for small molecules in water). Inserting typical values leads to characteristic times for the diffusion through the unstirred layer on the order of seconds. This is often long compared to the molecular processes immediately at the interface. In this respect, colloidal spheres (rather than planar substrates) are more favorable. Because the unstirred layer is never thicker than the sphere radius. The transport problem is an important consideration for sensor design.



General SPR Sensorgram – The target (pictured in blue - an antibody) is immobilized on the surface of a sensor chip. Then, the analyte (pictured in red - an antigen) is flowed over the surface and binding (Association or red curve) is seen. Once analyte is no longer supplied to the surface, the Dissociation of the analyte from the target is seen (blue curve). At the end of the dissociation, any remaining bound analyte is removed completely using a regeneration solution (green line). For simple 1:1 binding, the equilibrium dissociation constant (K_D), is defined as the dissociation rate constant (k_d) divided by the association rate constant (k_a).

Fig. 5.2
In the context of biointeraction analysis (BIA), one monitors the adsorption of an analyte as a function of time. In this process, equilibrium is reached rather rarely. The kinetics are analyzed, with the relevant fit containing the equilibrium constant as a parameter. One usually concludes with a "rinse". First, a desorption constant is determined. Furthermore, "unspecifically" adsorbed material is removed.
www.rdmag.com/article/2015/12/biggest-challenges-encountered-when-doing-spr-experiments
also: Reichert Technologies , 11/6/18

In biology, adsorption is often kinetically dominated in the sense that the molecules that reach a surface first irreversibly adsorb there, preventing adsorption of the molecules that arrive later. This is referred to as the "race for the surface." The fast molecules are usually the small molecules, because these can diffuse most quickly through the unstirred boundary layer.

5.2 The Gibbs adsorption isotherm and its consequences

Although the exact experimental determination of an adsorption isotherm is tedious, the isotherm (i.e., the adsorption equilibrium) is central to the discussion. We discuss three different adsorption isotherms, namely the Gibbs adsorption isotherm, the Langmuir isotherm, and the BET isotherm. The Gibbs adsorption isotherm has a certain special position here because it makes *no* assumptions about the interactions. It simply establishes a relationship between the amount adsorbed and the surface energy. It is also special in that on the right-hand side is not the activity of the adsorbed species in the bulk, but the derivative of the surface energy by the logarithm of that activity, $\delta\gamma/\delta(ln a_2)$. The relationship between adsorbed quantity (more precisely, surface excess) on the one hand and this differential ratio on the other is rigorously correct and universal. The derivation of the Gibbs adsorption isotherm is instructive. The relationship between $\delta\gamma/\delta(ln a_2)$ and a_2 depends on the molecular details. The Gibbs adsorption isotherm makes no statement about the latter relationship.

The Gibbs adsorption isotherm is based on a treatment of the interface within the framework of thermodynamics. Thereby, all extensive quantities are to be understood as excess quantities (Fig. 2.5). Excess quantities have already been introduced in the context of surface energy (Sect. 2.2). We recall a difficulty: the definition of the "Gibbs dividing surface" requires the existence of a majority component (a solvent).

When discussing the thermodynamics behind the surface energy, it was stated that:

Eq. 5.2.1

$$\mathrm{d}G_\sigma = \gamma \mathrm{d}A = \mathrm{d}H_\sigma - T\mathrm{d}S_\sigma$$

The subscript σ stands for the product of an area and an excess quantity in the Gibbs sense. In analogy to the fundamental theorem of thermodynamics we extend this equation to

Eq. 5.2.2

$$\mathrm{d}G_\sigma = \mathrm{d}H_\sigma - T\mathrm{d}S_\sigma + \sum_{Spezies} \mu_J \mathrm{d}n_{J,\sigma} = \gamma \mathrm{d}A + \sum_{Spezies} \mu_J \mathrm{d}n_{J,\sigma}$$

$n_{J,\sigma}$ is given here as $A\Gamma_J$ with J the index for the species. Note again the $\Gamma_1 = 0$. "1" denotes the majority component. The excess of the majority component is always zero according to the definition of Gibbs dividing surface. For a binary mixture we can write

Eq. 5.2.3

$$\mathrm{d}G_\sigma = \mathrm{d}H_\sigma - T\mathrm{d}S_\sigma + \sum_{Spezies} \mu_J \mathrm{d}n_{J,\sigma} = \gamma \mathrm{d}A + \mu_2 \mathrm{d}n_{2,\sigma}$$

At this point we use *homogeneity*. A function is "homogeneous"[40] in its arguments, if for each factor α applies

Eq. 5.2.4

$$f(\alpha x, \alpha y, \ldots) = \alpha f(x, y, \ldots)$$

The size G_σ is homogeneous in $n_{1,\sigma}$, $n_{2,\sigma}$ and A:

Eq. 5.2.5

$$G_\sigma(\alpha A, \alpha n_{1,\sigma}, \alpha n_{2,\sigma}) = \alpha G_\sigma(A, n_{1,\sigma}, n_{2,\sigma})$$

In prose terms, if you double an interface in size and all quantities of matter (while keeping the composition the same), the Gibbs Free Enthalpy of that system doubles. We derive Eq. 5.2.5 according to α:

Eq. 5.2.6

$$\frac{\partial G}{\partial(\alpha A)} A + \frac{\partial G}{\partial(\alpha n_{1,\sigma})} n_{1,\sigma} + \frac{\partial G}{\partial(\alpha n_{2,\sigma})} n_{2,\sigma} = G_\sigma(A, n_{1,\sigma}, n_{2,\sigma})$$

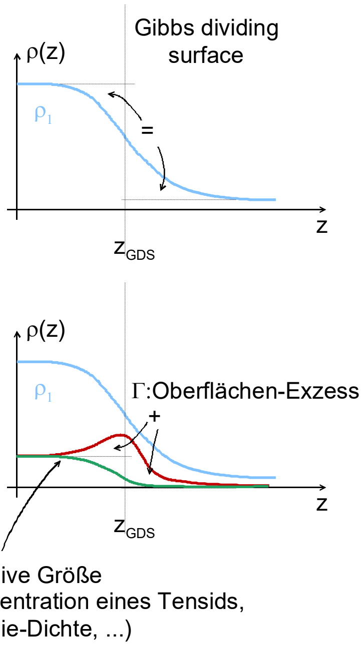


Fig. 5.3
On the definition of Gibbs-dividing surface and surface excess.

---

[40] It is also called "homogeneous of degree 1" because α occurs everywhere in the first power.

This relation is valid for all values of α, which includes $\alpha = 1$:

Eq. 5.2.7

$$\frac{\partial G}{\partial A}A+\frac{\partial G}{\partial n_{1,\sigma}}n_{1,\sigma}+\frac{\partial G}{\partial n_{2,\sigma}}n_{2,\sigma}=\gamma A+\mu_1 n_{1,\sigma}+\mu_2 n_{2,\sigma}=G_\sigma$$

The total differential, $\mathrm{d}G_\sigma$, is:

Eq. 5.2.8

$$\mathrm{d}G_\sigma=\gamma\mathrm{d}A+A\mathrm{d}\gamma+\mu_1\mathrm{d}n_{1,\sigma}+n_{1,\sigma}\mathrm{d}\mu_1+\mu_2\mathrm{d}n_{2,\sigma}+n_{2,\sigma}\mathrm{d}\mu_2$$

We again use that the surface excess of component 1 is zero and subtract Eq. 5.2.3 from Eq. 5.2.8

Eq. 5.2.9

$$\begin{aligned}0 &= A\mathrm{d}\gamma+n_{2,\sigma}\mathrm{d}\mu_2 = A\mathrm{d}\gamma+A\Gamma_2\mathrm{d}\mu_2\\ &= A\mathrm{d}\gamma+A\Gamma_2\mathrm{d}\left(\mu_2^{\varnothing}+RT\ln a_2\right)\\ &= A\mathrm{d}\gamma+RTA\Gamma_2\mathrm{d}\left(\ln a_2\right)\end{aligned}$$

Eq. 5.2.9 is a variant of the *Gibbs-Duhem equation*. In step 2, it was used that $n_{2,\sigma} = A\Gamma_2$. Step 3 made use of the definition of activity in terms of chemical potential. The area *A* cancels. We rearrange:

Eq. 5.2.10

$$\Gamma_2=-\frac{1}{RT}\left(\frac{\partial\gamma}{\partial\ln a_2}\right)_T\approx-\frac{1}{RT}\left(\frac{\partial\gamma}{\partial\ln\left(c_2/c_2^{\varnothing}\right)}\right)_T$$

This is the *Gibbs adsorption* isotherm. In step 2, the activity coefficient was assumed to be unity.[41]

Before discussing the Gibbs adsorption isotherm further, we briefly address the question of what happens when a majority component does not exist. (A majority component was required in the definition of the Gibbs dividing surface, GDS). Basically, one then falls back to the situation that we also found in the context of thermodynamic principles. Then the GDS has to be placed in such a way *that* the Gibbs adsorption isotherm holds. One wonders, of course, what then is the statement of the Gibbs adsorption isotherm. It is then a definition of the surface excess. If there is the majority component, one can measure Γ_2 in principle. The emphasis here is on in principle, because such measurements are complicated at interfaces between two fluids. (The fact that the position of the GDS as postulated by Gibbs indeed produces the correct surface excess has not been proved here, by the way).

If there is no majority component, the Gibbs adsorption isotherm must be reformulated in two other places. First, there is then a surface excess of component 1. Second, the chemical potentials μ_1 and μ_2 are not independent. For binary mixtures, the Gibbs-Duhem relation applied to the bulk leads to the relation

Eq. 5.2.11

$$0=n_1\mathrm{d}\mu_1+n_2\mathrm{d}\mu_2$$

---

[41] Remember: $a = \gamma c/c^{\ominus}$ with γ the activity coefficient (not the surface energy) and $c^{\ominus}$ the standard concentration.

The extension of Eq. 5.2.9 to systems without high dilution is therefore.

Eq. 5.2.12

$$\begin{aligned} 0 &= A\mathrm{d}\gamma + n_{1,\sigma}\mathrm{d}\mu_1 + n_{2,\sigma}\mathrm{d}\mu_2 \\ &= A\mathrm{d}\gamma + A\Gamma_1\mathrm{d}\mu_1 + A\Gamma_2\mathrm{d}\mu_2 \\ &= A\left(\mathrm{d}\gamma + \left(-\Gamma_1\frac{n_2}{n_1} + \Gamma_2\right)\mathrm{d}\mu_2\right) \end{aligned}$$

It follows that

Eq. 5.2.13

$$\Gamma_2 - \frac{n_2}{n_1}\Gamma_1 = -\frac{1}{RT}\left(\frac{\partial\gamma}{\partial\ln a_2}\right)_T$$

At high dilution, the second term on the left-hand side can be neglected, which recovers the standard Gibbs adsorption isotherm.

We discuss two application examples. Fig. 5.4 shows the surface tension of water versus the logarithm of the concentration of the surfactant CTAB (cetyl trimethylammonium bromide). Note the logarithmic x-scale. If we approximate the activity coefficient to be 1, the concentration is proportional to the activity. Because the *x*-axis in this graph was chosen to be logarithmic, the slope is equal to $-RT\Gamma_2$.

The kink in Fig. 5.4 is of minor interest here. The kink occurs at the critical micelle concentration (CMC). We will return to micelles in Sect. 7. At concentrations larger than the CMC, the chemical potential of the CTAB molecules in the water phase does not increase further because the excess molecules are trapped in the micelles (Fig. 5.6).

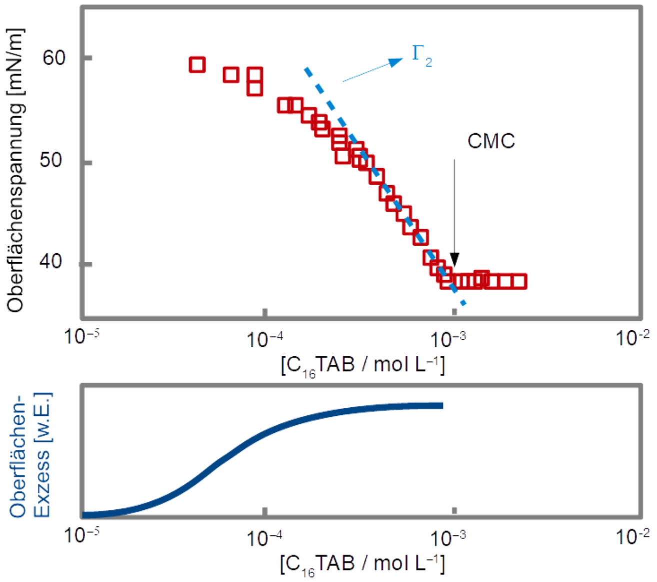


Fig. 5.4
The surface tension of an electrolyte solution decreases when a surfactant is added to the subphase. Note the logarithmic scale on the *x-axis*. From the slope, calculate the surface excess of the surfactant. The kink occurs at the critical micelle formation concentration (CMC).
Adapted from .
Goodwin, J., Colloids and Interfaces with Surfactants and Polymers.

At $c < c_{CMC}$ the curve $\gamma(\ln c_2)$ shows downward curvature, meaning that Γ_2 increases with increasing concentration, as one would expect. Remarkably, there is a region below the CMC (about half a decade wide), in which the slope is nearly constant, meaning that Γ_2 is almost independent of the surfactant concentration in the water phase. In this range, the surfactant is in a particularly stable conformation, which causes the packing density to stay almost constant, even when the chemical potential of CTAB varies. This favorable packing is sketched in Fig. 5.5. The alkyl chains are largely stretched. There is an attractive interaction between the stretched chains that requires the stretched chains to be close to each other. This requirement brings about the wide range with $\Gamma_2 \approx const$. This brings us to the important fact that surfactants are characterized distinguished by *two* separate properties. The first is amphiphilicity. Surfactants are hydrophilic at one end and hydrophobic at the other. (“lipophilic" here is synomnymous to “hydrophobic”.) Ethanol is also amphiphilic, but is not considered a surfactant. The second property of surfactants are lateral interactions and the ability to self-assemble (Sect. 7). Amphiphilicity is only one of two prerequisites for the formation of micelles, organized layers at the water/air interface, and lamellae.

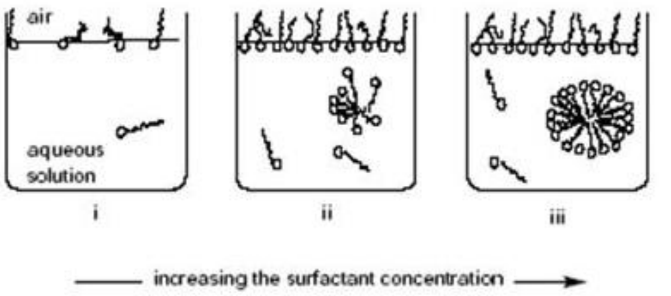


Fig. 5.5
Arrangement of surfactant molecules in aqueous solutions: (i) at low concentrations – some surfactant molecules adsorb at the air-water interface; (ii) at higher concentration a monolayer is formed at the air-water interface; (iii) an increase of surfactant concentration does not change the state of this monolayer. Surface. All additional surfactant molecules aggregate as micelles.
www.tau.ac.il/~phchlab/experiments_new/surface_tenstion/theory.html

The surface excess can also be negative. Molecules can be *depleted* at the interfaces. Many salts are depleted at the air-water interface, which has a simple explanation. According to the Debye-Hückel theory, each ion is surrounded by a counter ion cloud.[42] Because the counterion cloud is cut off from the interface, many ions avoid the interface. $\Gamma_2 = -RT\,\partial\gamma/\partial(\ln a_2)$ is negative and the surface energy *increases* with increasing salt concentration (Fig. 5.6).

Similar to most results of the Debye-Hückel theory, there are counterexamples. Regardless of the counter-ion cloud (meaning: the loss of half of it), ions sometimes are enriched at the air-water interfaces. This is one of the "ion-specific effects" (ISEs), related to the Hofmeister series (Sect. 10.8). The consequences of this enrichment are known from thunderstorms. Clouds are charged electrically, when heavy hailstones (descending) collide with light hailstones (ascending). Triboelectricity transfers charge between them. (Triboelectricity sometimes occurs between shoes and carpets.) The triboelectricity occurs because there are ions on the surface of the grains. The details are complicated and variable: not all thunderstorms are the same.

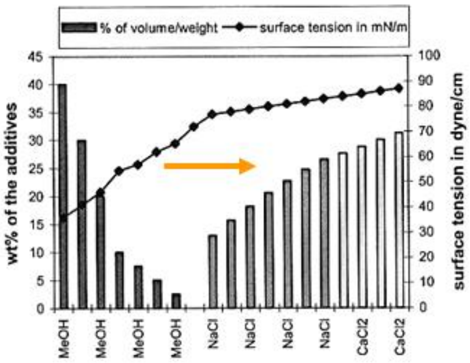


Fig. 5.6
By adding methanol, NaCl, or $CaCl_2$, one can control the surface tension of aqueous solutions within certain limits. Addition of salt can increase the surface energy, because salts are often deposited at the interface.
Source: R&D, Sartorius AG

5.3 The Langmuir isotherm and the BET isotherm

Unlike the Gibbs adsorption isotherm, the Langmuir isotherm makes specific assumptions about the interactions between the surface and the adsorbate. These are:

- There is a fixed number of adsorption sites. These can either be empty or occupied once. Multi-layer adsorption is not permitted.
- All adsorption sites are equivalent.
- Apart from double occupation not being allowed, there are no interactions between the adsorbed molecules. .
- The particles do not change their properties upon adsorption. (This is often not true for proteins. These often "denature" upon adsorption and then are no longer biologically active.)

---

[42] The counterion cloud does not contain many counter ions. That would violate electroneutrality. The counterion cloud is an increased probability density for counterions.

For this situation, one might, in principle, calculate a free energy of the adsorbate per unit area (equivalent to γ), take the derivative to $\ln(a_2)$, and obtain Langmuir isotherm from the Gibbs adsorption isotherm. However, another argument turns out to be easier. Let θ be the fraction of occupied sites (the "coverage)". At equilibrium, θ is constant in time. The desorption rate is proportional to θ. Name the rate constant k_{des}. The adsorption rate is proportional to the fraction of vacant sites (to $1-\theta$) and to the pressure of the gas (or the concentration of the species of interest in the bulk). In the following, we use the pressure p. Name the kinetic coefficient k_{ads}. In equilibrium both rate are equal and opposite

Eq. 5.3.1

$$0=\dot{\theta}=-k_{des}\theta+-k_{ads}\left(1-\theta\right)p$$

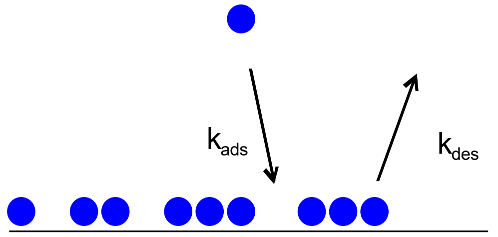


Fig. 5.7
One can derive the Langmuir isotherm from a kinetic model. In this model, the binding constant K is the ratio of the velocity constants k_{ads} and k_{des} .
ms.zneb.at/html/chemistry1/chse28.htm

Solving for θ leads to

Eq. 5.3.2

$$\theta=\frac{k_{ads}\,p}{k_{des}+k_{ads}\,p}=\frac{p}{K^{-1}+p}$$

In the second step, the ratio k_{des}/k_{ads} was renamed as K^{-1}. One might have called this ratio K. It is called K^{-1} because K as defined this way can be viewed as binding constant, also known from chemical reactions ([AB] = K [A][B] with AB the complex of A and B and […] the concentration in units of mol/L[43]). The binding constant is given as k_{assoc}/k_{diss} with k_{assoc} and k_{diss} the kinetic coefficients for association and dissociation.

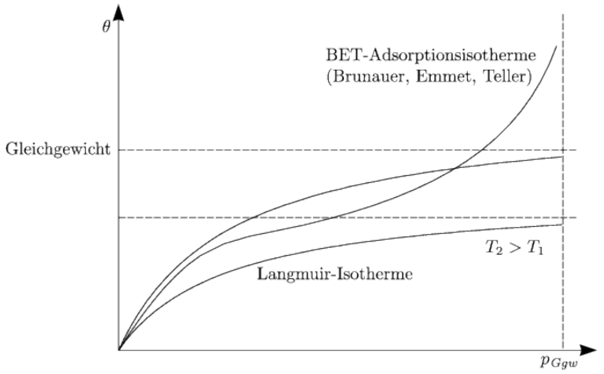


Fig. 5.8
Langmuir isotherm and BET isotherm.
ms.zneb.at/html/chemistry1/chse28.htm

Note: In liquids adsorption is *competitive adsorption*. The adsorption of a solute molecules involves the displacement of solvent molecules. The same applies to the effective van-der-Waals forces between molecules in liquids. These are always competing interactions. This is sometimes called and "Archimedes' principle".

Usually, the coverage itself is not experimentally accessible, but only the adsorbed amount n_{ads}. We write $\theta = n_{ads}/n_{sat}$ with n_{sat} the adsorbed amount at saturation. We us Eq. 5.3.2 and write

Eq. 5.3.3

$$n_{ads}=\frac{p}{K^{-1}+p}n_{sat}$$

---

[43] Note the similarity in the shape of the Langmuir isotherm and the Michaelis-Menten kinetics. This similarity has its roots in saturation.

The right-hand side of this equation contains three free parameters. All these can be measured and may be interesting in different situations. Before we come to these situations, some further remarks:

- In solution, the pressure p is replaced by the concentration c.
- As always, one should in principle replace pressure and concentration with fugacity and activity because there may be interactions between the molecules in the bulk phase, which will affect the adsorption equilibrium.
- The central characteristic of the Langmuir isotherm is *saturation* (Fig. 5.8). If a particular species of molecules were to attach to a surface consisting of the *same* molecules, then those molecules would also cluster in the bulk and would not be well soluble, in consequence. Those molecules, which are well soluble, at the same time do not form multilayers at surfaces.[44]

  From this last observation, a strategy for the preparation of low-fouling surfaces in biotechnology and medical technology is derived. The primary process in biofilm formation is the adsorption of proteins. If a particular species of proteins dominates, a dense, smooth packing of proteins of this same species can prevent further adsorption of proteins.

Sometimes the saturation in the sense of the Langmuir isotherm is not perfect in the sense that multilayer formation occurs at high bulk concentrations. This possibility is accounted for by the *BET isotherm* ("BET" for to Brunauer, Emmet, and Teller). A BET isotherm is shown in Fig. 5.8. Note: The BET isotherm also shows a plateau caused by (incomplete) saturation.

The algebraic form proposed by Brunauer, Emmet and Teller is

Eq. 5.3.4

$$\theta = \frac{n}{n_{sat}} = \frac{\tilde{K}\,\dfrac{p}{p_{sat}}}{\left(1-\dfrac{p}{p_{sat}}\right)\cdot\left(1+\dfrac{(K-1)\,p}{p_{sat}}\right)}$$

The BET isotherm has a microscopic justification, but it may also be viewed as heuristic.[45] Note: $\tilde{K}$ is dimensionless in this equation. In the limiting case of $p_{ads} \rightarrow 0$, θ is proportional to $\tilde{K}p/p_{sat}$, not to Kp, as in the Langmuir isotherm. One could, of course, replace $\tilde{K}$ by K/p_{sat}, but we stick to the usual form.

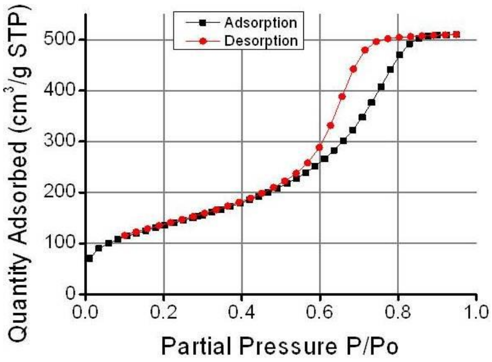


Fig. 5.9
A data set as it appears in BET porosimetry. The plateau corresponds to the saturation. From the amount of substance taken up by the porous medium on the plateau, the integral surface area of the pore system is inferred
www-omcs.materials.ox.ac.uk/facilities/surface-analysis/bet-surface-area-analysis.html.

Fig. 5.9 shows data obtained with *BET porosimetry*. Essential for BET porosimetry is also the saturation. One can see

---

[44] When gases adsorb to a surface from the vapor phase, wetting improves as the critical point is approached. “Wetting” here denotes the tendency of adsorbed films to diverge in thickness at $p \rightarrow p_{sat}$. At the same time, it denotes the spreading of droplets at $T \rightarrow T_{crit}$. This type of “critical wetting” goes back to clusters being formed at $T \rightarrow T_{crit}$. Critical wetting would be technically interesting if it were robust. Unfortunately, it is not.

[45] The term "heuristic" is commonly used in the sense that the argument in question is formulated in limited time and with limited effort, even though a deeper understanding may be lacking. Many empirical equations are heuristic in this sense. From a methodological point of view, the term can be sharpened somewhat. Heuristic arguments are part of the process of understanding. (Archimedes shouted "Eureka" as he understood buoyancy while sitting in a water tub). A heuristic equation describes experimental facts without the equation being grounded in the first principles of physics. The latter may follow as a second step. The Planck formula for black body radiation was heuristic in this sense. It was a guess, which was later connected to the fundamental principles of physics by deriving this same formula from the postulate of quantized radiation. In the process, the Planck formula changed the principles of physics. A guess here lead to understanding.

from the data at what absorbed amounts saturation occurred. With an assumption for the area per molecule, one can calculate the internal surface area of a porous medium from this. One writes

Eq. 5.3.5

$$A_{pore} = \bar{A}_{ads} n_{ads} = N_A A_{molecule} n_{ads}$$

A_{pore} is the inner surface of the porous medium (i.e. the surface accessible to the adsorbate). $\bar{A}_{ads}$ is area per mole of the adsorbate, n_{ads} is the adsorbed amount. At this point it is easier to use the area per molecule ($A_{Molecule} \approx 10$ Å^2). Either way, the Avogadro constant N_A, shows up somewhere.

Talking of porous media: There may be capillary condensation. Capillary condensation can explain the rise on the right-hand side of the BET isotherm without the need to postulate multilayers.

5.4 Adsorption energies

The Gibbs adsorption isotherm expresses in algebraic form that adsorption is driven by a decrease of the surface energy. Furthermore, the Gibbs adsorption isotherm can be used to convert from Γ_2 (adsorbed amount) to $\partial\gamma/\partial(\ln(a_2))$ and vice versa. To obtain $\gamma(a_2)$, one would have to integrate $\partial\gamma/\partial(\ln(a_2))$ over a_2, which of course requires accurate measurements. The advantage: This method also works on solids or powders. Nevertheless, this procedure it is not widely applied. In more general terms, an adsorption isotherm may have structure (see e.g. Fig. 5.10) and the structures can be attributed to a dependence of the adsorption energy on the degree of coverage.

There is direct access to the energy of adsorption with calorimetry. Since the adsorbed amount is only a few µmol/cm^2, such measurements require good sensitivity. Using a thin membrane as the substrate helps. Thin membranes have a small heat capacity, which is why even small amounts of heat cause as measurable increase in temperature. One can also increase the surface area by replacing the substrate with a large number of colloidal spheres. In both cases, it helps if adsorption is fast because there are instrumental drifts. Fast adsorption requires fast exchange between the surface and the bulk and/or fast mixing.

Related to calorimetry is thermal desorption spectroscopy (TDS). A sample is heated in vacuum and the desorbed molecules are fed to a mass spectrometer (Fig. 5.11).

5.5 Structure and dynamics of adsorbates

In addition to the energetics, one is always interested in the structure of the adsorbate, possibly also in its microscopic dynamics. For adsorbates on single crystal surfaces in high vacuum, there have been spectacular successes in this regard in recent decades (2007: Nobel Prize to Gerhard Ertl, Berlin). An important application is heterogeneous catalysis in gases. A molecularly detailed understanding - for model systems - has also been achieved in an electrochemical context. Methods of investigation are the tools of electrochemistry (especially current-voltage relations) and the scanning tunneling microscope.

For many technical adsorbates, ab-initio modeling is out of reach. This is especially true for structured surfaces. Catalyst grains (often metallic nanoparticles, themselves heterogeneous with co-catalyst) are ceramic-supported, so not even the structure is well known. Research on heterogeneous catalysis is quite largely empirical. (The same is true for the use of surfactant mixtures for colloid stabilization).

Linear alkane chains with thiol end groups on gold form a special case. These molecules are laterally mobile. They reorganize *after* adsorption and form an ordered structure (Fig. 5.12). This is referred to as a "self-assembled monolayer" (SAM). The structure is particularly well ordered on the Au111 surface. The structures are robust in the sense that they are preserved when a few molecules do not consist of alkane chains. The alkane chains then form a matrix. Quite often, a subgroup of molecules contains groups at the end which can undergo chemical reactions (amines, carboxyl groups, ...) or which are specifically recognized (biotin, recognized by streptavidin). SAMs are called a "platform" for surface functionalization.

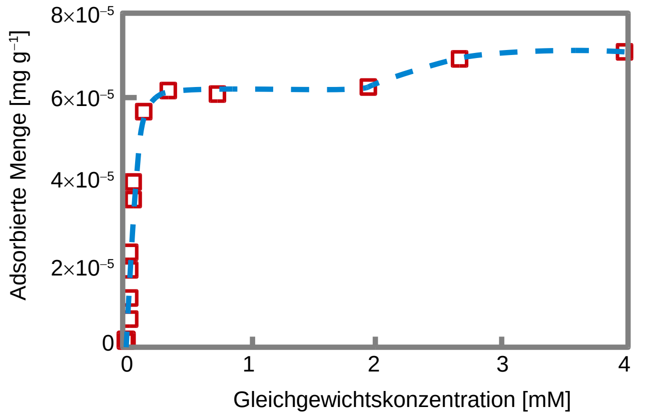


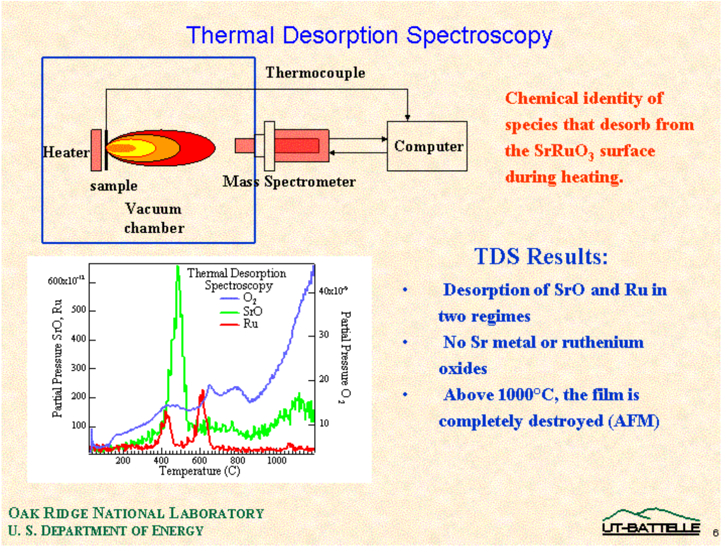
Fig. 5.10
An adsorption isotherm may have a non-trivial structure. The second stage is the result of reorganization of the adsorbate at high coverage. The surfactant CTAB was adsorbed. The substrate consists of kaolin particles (a clay mineral also used in porcelain production).
Adapted from . Goodwin, J., Colloids and Interfaces with Surfactants and Polymers.

5.6 Inverse Gas Chromatography

iGC also gives access to the surface energies. The longer the retention time, the stronger the adsorption.

<u>Remember</u>

- Adsorption occurs to the extent that it lowers the surface energy.
- On solid surfaces, adsorption equilibrium is often not achieved.
- The Gibbs adsorption isotherm states that the surface excess Γ_2 follows the relation $\Gamma_2 = -RT\, \partial\gamma/\partial \ln(a_2)$.



Thermal Desorption Spectroscopy

Chemical identity of species that desorb from the $SrRuO_3$ surface during heating.

TDS Results:

- **Desorption of SrO and Ru in two regimes**
- **No Sr metal or ruthenium oxides**
- **Above 1000°C, the film is completely destroyed (AFM)**

OAK RIDGE NATIONAL LABORATORY
U. S. DEPARTMENT OF ENERGY

UT-BATTELLE 6

Fig. 5.11
About Thermal Desorption Spectroscopy (TDS)
wiki.ornl.gov/sites/cnms_imaging/website/nanotransport/stabilityofsrru03.html.

- The Langmuir isotherm and the BET isotherm are characterized by *saturation*.
- Surfactants are amphiphilic and they self-organize due to interactions between the hydrophobic alkyl chains.
- The energy of adsorption can also be determined calorimetrically.
- The structure and dynamics of adsorbates are the subject of current research.
- Alkane thiols on gold form self-assembled monolayers (SAMs).

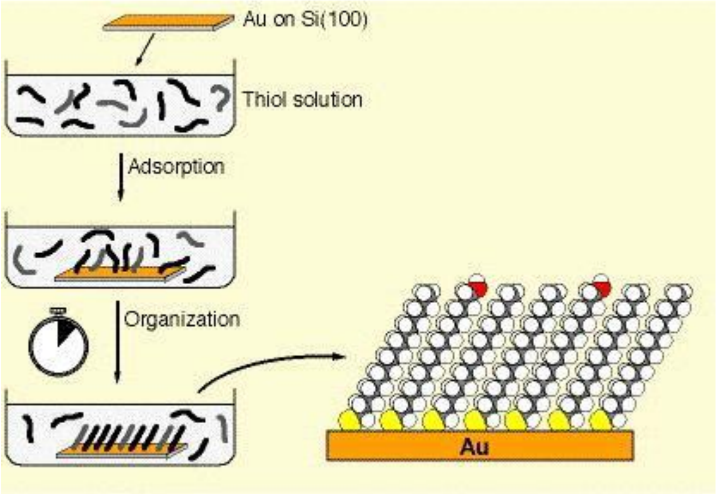


Fig. 5.12
On the formation of a "self-assembled monolayer" on a gold substrate by thiolation.
www.ifm.liu.se/applphys/molphys/research/self/

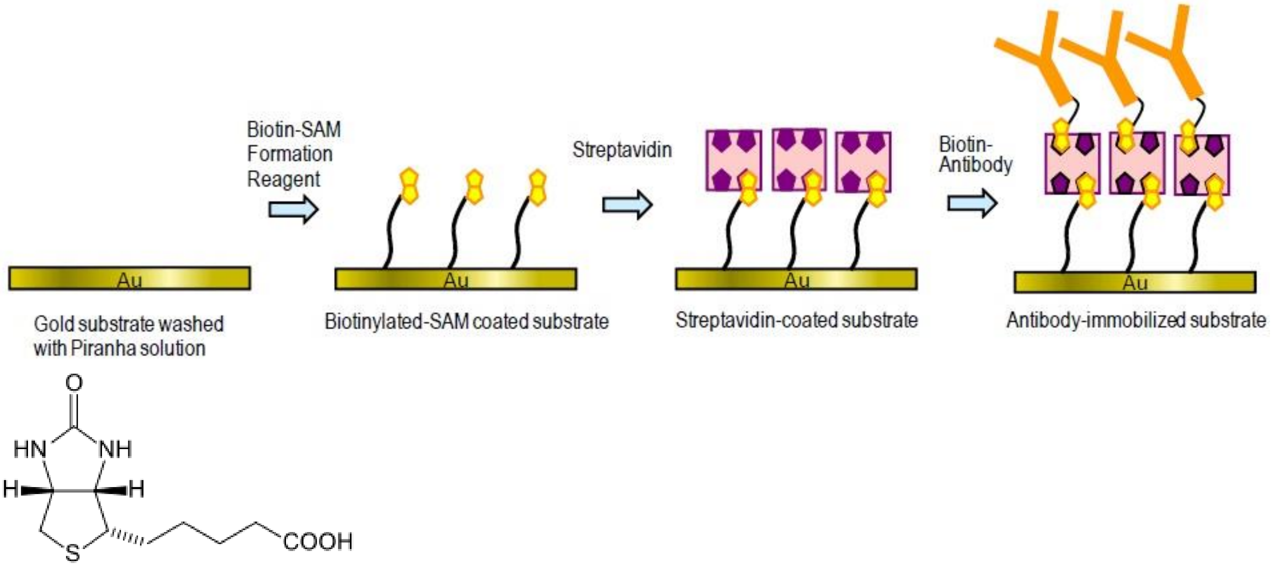


Fig. 5.13
Top: Immobilization of antibodies on gold via a thiol layer, biotin, streptavidin and biotin.
www.dojindo.com/store/p/797-Biotin-SAM-Formation-Reagent.html
Bottom: Biotin
en.wikipedia.org/wiki/Biotin

6 Wetting and contact angle

6.1 General

According to the BET isotherm (Sect. 5.3), the thickness of an adsorbate layer continuously approaches infinity as the vapor pressure approaches the saturation vapor pressure. This is realistic for porous media because of capillary condensation. For planar surfaces it is rather the exception. Such continuous divergence leads to complete wetting, and complete wetting of a surface by a liquid is not often found. As shown in Fig. 6.1 illustrated, at the edge of a droplet the vapor pressure is always close to the saturation vapor pressure. Then liquid would have to condense onto the surface in such a way that macroscopic thicknesses are reached. This amounts to spreading of the droplet.

Even though this complete wetting is not often observed, there is always a microscopically thin, "prewetting film" in the vicinity of a droplet. Fig. 6.2 shows a scanning electron micrograph of a cooled glass drop. (Such images cannot be obtained in liquid form because most liquids evaporate in vacuum). This prewetting film is not usually seen by eye. Nevertheless, one must always keep in mind the possibility of its existence. This is especially true for contamination of clean-apparent slides that one has touched with bare fingers. The contamination with a molecular monolayer of fats spreads over the slide after touching it, without being seen. One cyn visualize the contamination by wetting the substrate with water and blowing the water off an air gun. This is a simple test for cleanliness.

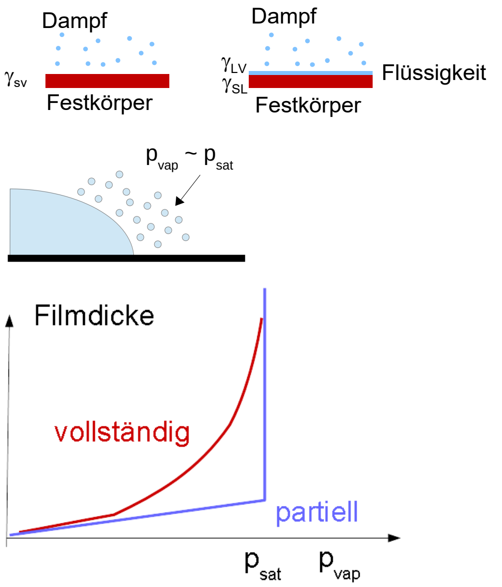


Fig. 6.1
When the thickness of an adsorbate in the limit of $p_{vap} \rightarrow p_{sat}$ tends towards ∞, droplets spread out because p_{vap} is about equal to p_{sat} at the edge of the drop.

The thickness of the prewetting film is of the order of the range of the intermolecular forces. This range amounts to a few molecular diameters. Inside the prewetting film, all molecules feel both interfaces. The prewetting film not a bulk phase. Prewetting implies that there is an attractive interaction between the substrate/film interface and film/vapor interfaxe. The attractive interaction keeps the film thickness finite.

Of some special relevance are prewetting films of air in water on hydrophobic surfaces. One can see these films ("nanobubbles", "nanopancakes") with the AFM. IR spectroscopy proved that the features seen with the AFM indeed contained a gas. A rotational vibrational spectrum was seen.[46] These thin layers of air facilitate the slip of the water on that surface (section 9.1).

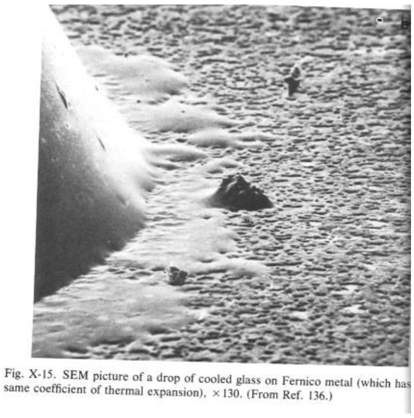


Fig. X-15. SEM picture of a drop of cooled glass on Fernico metal (which has same coefficient of thermal expansion), ×130. (From Ref. 136.)

Fig. 6.2
A prewetting film
In this case, it is a solidified glass, imaged with a scanning electron microscope.
Adamson, A. W.; Gast, A. P., Physical Chemistry of Surfaces, John Wiley: 1997.

As a rule, droplets do not spread completely, but form a finite contact angle with the surface. Consider the drops sketched in Fig. 6.4. The contact angle is calculated from an equilibrium of forces. A surface

---

[46] Zhang, X. H.; Quinn, A.; Ducker, W. A., Nanobubbles at the interface between water and a hydrophobic solid. *Langmuir* **2008,** 24, (9), 4756.

tension is a force per length. This length is perpendicular to the plane of the paper. The substrate is planar and solid. The meniscus can therefore only move in the substrate plane. This is a constraint condition. The force exerted by the liquid-vapor interface acts obliquely to the substrate plane. Only the part of this force that acts parallel to the interface is included in the force balance. The force equilibrium is

Eq. 6.1.1

$$\gamma_{SV} = \gamma_{SL} + \gamma_{LV} \cos\theta$$

This is the *Young equation*. The indices *S*, *L* and *V* denote the "solid", the "liquid" and the "vapor". Solving for cos θ results is

Eq. 6.1.2

$$\cos\theta = \frac{\gamma_{SV} - \gamma_{SL}}{\gamma_{LV}}$$

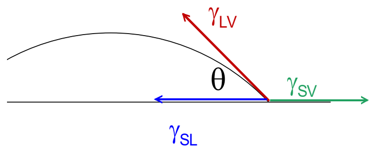


Fig. 6.4
The contact angle is calculated from an equilibrium of forces (Eq. 6.1.1).

Thus, wetting occurs when the sum of the surface energies $\gamma_{SL} + \gamma_{LV}$ is less than the surface energy of the bare solid, γ_{SV}.[47] A side remark: It is difficult to test the validitiy of the Young equation experimentally because the equations contains surface energies of solids.

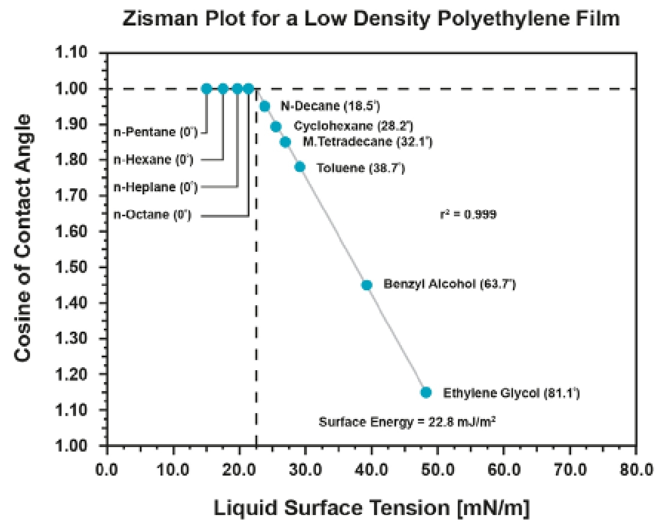


Fig. 6.3
Zisman plot for a polyethylene surface with a series of liquid. *www.biolinscientific.com/zafepress.php_pdf_Attension_Theory_Note_FAT_TN_4_sfe.pdf*

Note: The contact angle is independent of what the conditions are at the molecular level at the three-phase boundary.

Unfortunately, it is not possible to calculate the two surface energies γ_{SV} and γ_{SL} separately from the contact angle, because the right-hand side of Eq. 6.1.2 contains three unknowns, while there are only two parameters determined in experiment, which are θ and γ_{LV}. Still: the higher the surface energy of a solid, the smaller the contact angles with liquids. Zisman has turned this argument into recipe. Fig. 6.3 shows a Zisman plot. The y-axis shows cosine of the contact angles for a series of homologous liquids. These liquids might be alkane chains or alcohols with different chain lengths. It is found that cosθ increases as γ_{LV} decreases. Plotting cos θ against γ_{LV} leads to a straight line, which can be extrapolated to cos θ = 1. Zisman calls the value on the x-axis γ_c. γ_c is a related to the surface energy of the solid substrate. Polyethylene is wetted by nonpolar fluids with low surface energy, but not by fluids with higher surface energy.

---

[47] One might expect that the sum of two surface energies would often be larger than the surface energy of a single third interface. The opposite is predicted by simple models of nonpolar substances. In this picture, the surface energy between substances A and B is written as $\gamma_{AB} \approx \alpha(\delta_A - \delta_B)^2$ with α a prefactor that brings about the correct dimension, and δ a dimensionless measure of internal cohesion. Use $\delta_{solid} = 2$, $\delta_{liq} = 1$ and $\delta_{vap} = 0$. Then $(\delta_{solid} - \delta_{vap})^2 = 4$, while $(\delta_{solid} - \delta_{liq})^2 + (\delta_{liq} - \delta_{vap})^2 = 2$. Following this argument, complete wetting should be common. In practice, however, it is rather rare. Dirt plays a role, because solid surfaces with high energy are rarely clean.

The considerations leading to the Zisman plot are complicated. In all cases, there is an underlying - more or less realistic - assumption as to how the surface energy γ_{SL} depends on the surface energies γ_{SV} and γ_{LV} . With such a relation one reduces the number of unknowns in the Young equation (Eq. 6.1.1) by one and then one can invert this equation. One can understand such a relation as a combining rule.[48] The combining rules make assumptions about how the interactions between two media (here: solid and liquid) can be described simply, if the interactions of the molecules concerned with the same molecules are known (the latter here inferred from the surface energies with the vapor phase). The mixing rules are manifold and all of them are of limited viability.

Since the assumptions underlying the Zisman plot are quite broad, it is not surprising that homologous series of polar or nonpolar fluids do not lead to the same value of γ_c . One then speaks of γ_p and γ_d . ("*p*" for polar and "*d*" for dispersive, i.e. nonpolar). Sometimes other homologous series and other parameters are introduced. The procedure is semiempirical and oriented to practical requirements.

In deriving Young's equation, it was assumed that the meniscus was free to move on the surface. Because it usually is not, there is a difference between the advancing contact angle and the receding contact angle. The difference between the two angles ($\Delta\theta$, a measure of contact angle hysteresis) provides information about the roughness of the surface and about chemical heterogeneities (dirt).

When determining the contact angle hysteresis, make sure that the menisci have come to a standstill before the measurement. The contact angles during movement are part of the dynamic wetting. Dynamic wetting and dewetting are complicated and at the same time technically important. The difficulties stem from the fact that molecular scale processes at the three-phase boundary line play a role. (In equilibrium one does not need to consider them. One can consider the forces in Fig. 6.4 far away from the contact line).

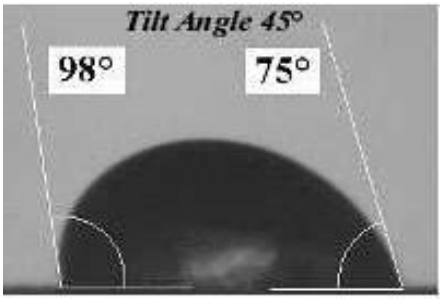


Fig. 6.5
On contact angle hysteresis
www.pocketgoniometer.com/main.php?cont=hysteresis&lang=en

For wetting on rough surfaces (Fig. 6.6), a distinction is made between two limiting cases, named after Wenzel and after Cassie and Baxter. In the former case, the liquid penetrates into the troughs. One calculates an apparent macroscopic surface energy between solid and liquid. This increases with increasing roughness because the true surface area increases. Therefore, in this limiting case, the contact angle decreases with increasing roughness. In the limiting case named after Cassie and Baxter, the liquid wets only the tips. The apparent macroscopic contact angle increases. These two limiting cases also play a role in the context of pressure sensitive adhesives (PSAs, an example being Tesa film). If the adhesive fully wets a rough surface (i.e., it is soft enough to penetrate the indentations), roughness improves adhesion. Otherwise, it counteracts adhesion.

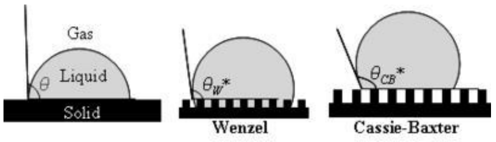


Fig. 6.6
Wetting on rough surfaces
www.tau.ac.il/~phchlab/experiments_new/surface_tenstion/theory.html

The Cassie-Baxter case is the basis of ultrahydrophobic surfaces and the lotus effect. Ultrahydrophobic surfaces are both hydrophobic and rough. Often, "fractal" roughness is required, i.e., a surface that is rough at all scales. Ultrahydrophobic surfaces are technically attractive because all droplets roll off, taking dirt with them. These

---

[48] en.wikipedia.org/wiki/Combining_rules. An example for a combining rule is $\gamma_{SL} \approx (\gamma_{SV}\,\gamma_{LV})^{1/2}$

surfaces have therefore been called "self-cleaning." The technical breakthrough has failed to materialize. For one thing, these surfaces are often mechanically fragile. Another problem is dirt, which - firstly - settles in the valleys (thus reducing the roughness) and which - secondly - turns the hydrophobic surface into a surface with a contact angle in the middle range.

Today, the industry speaks of "easy-to-clean" rather than "self-cleaning". Among other things, "easy-to-clean" means scratch-resistant and in the sense that graffiti colors can be tackled "with a root brush" without destroying the paint.

"Superhydrophilic" surfaces lead to droplet spreading. This would be attractive for anti-fog applications. However, very hydrophilic surfaces are unsuitable for this purpose because contamination quickly reduces the hydrophilicity. One uses swellable coatings. The water penetrates a hydrophilic matrix, spreads there, and reappears on the surface next to the droplet, causing the spreading. With this mechanism, contamination is less of a problem because the contaminants must penetrate the interior of the film to reduce the antifog effect (Fig. 6.7).

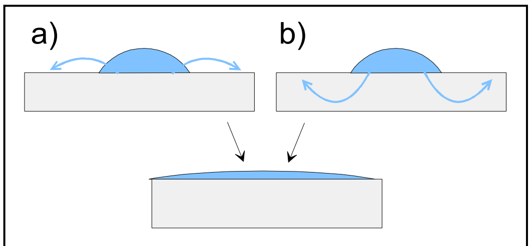


Fig. 6.7
Mechanisms of droplet spreading.
a) Spreading on the surface
b) Transport inside the film.
An antifog layer based on spreading in the film is less susceptible to dirt and thus more durable than an antifog layer based on a hydrophilic surface.

The wetting of a surface by an electrolyte can be switched electrically ("electrowetting", section 2.4).

Wettability plays a major role in offset printing.

Wetting and dewetting processes in porous media are extraordinarily complicated. Soil hydrology deals with this question. Mechanics, chemistry, capillarity and biology are mixed here.

Remember

- Complete wetting and complete dewetting are rare. This has to do with saturation in the context of adsorption.
- At the microscopic level, there is often a prewetting film.
- The three-phase boundary line is often anchored (roughness, chemical heterogeneities). The contact angle hysteresis provides information about such heterogeneities.
- If the three-phase boundary line is not anchored, there is an equilibrium of forces. The Young equation applies: $\gamma_{SV} = \gamma_{SL} + \gamma_{LV} \cos\theta$.
- Using the Zisman plot, one can infer the surface energy of solids from contact angles. However, the concept makes assumptions.
- Ultrahydrophobic surfaces are rough surfaces.

7 Surfactants and self-organization

7.1 Mechanisms of self-organization

Self-organization has already been briefly discussed in Sect. 5.2 in connection with the formation of micelles. What is "special" about surfactants is their ability to self-organize. In the following, self-organization or self-assembly means the formation of spatially limited structures. The structures in question should be equilibrium structures. Robust structures are in view. This means that the structure formation principles do not depend critically on molecular details. Self-organization in this sense is central to the emergence of complex systems, including life.

The following objects do not count as self-organized structures:

a) Crystals
Crystals are not limited in size.

b) Objects created by terminated growth.
Two examples:
- Gold nanoparticles (AuNPs) are formed from gold acid ($H^+(AuCl_4)^-$) in the presence of citrate, where the citrate acts as a reducing agent. In a second step, the citrate adsorbs to the surfaces of the growing particles and passivates them against further growth. After citrate is loaded, it stabilizes the AuNPs against aggregation. A size of about 10 nm is easily predictable.
- Filter paper (Fig. 7.1) is usually prepared by spinodal segregation of a polymer and a solvent in a precipitation bath. The structures are initially filigree and then coarsen. The process comes to an end by solidification of the polymer phase. The timing of the solidification determines the pore size.

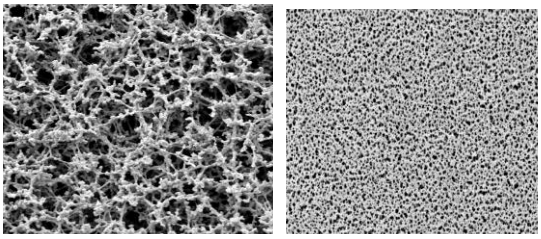


Fig. 7.1
Various membranes produced by precipitation of a polymer phase from a solvent.
The sizes are 100 µm (left) and 40 µm (right).
Provided by Dr. Volkmar Thom, Sartorius AG, Goettingen, Germany

c) Dissipative structures
Far away from thermal equilibrium, spatiotemporal patterns sometimes form, that are similarly important to self-organization in the formation of complex systems. Examples are the convection rolls in Rayleigh-Benard convection, the concentration gradients (or waves) that arise as solutions of nonlinear, autocatalytic reaction-diffusion equations (Belousov-Zhabotinsky reaction), the heartbeat, embryogenesis, and brain activity.

The following text is limited equilibrium structures. Essential mechanisms are:

a) The frustrated incompatibility
b) Topological constraints and in particular a high energetic price for edges of lamellas.
c) Spontaneous curvature of lamellae
d) Helicity
e) Short-range attraction combined with long-range repulsion (SALR).

f) Cooperative specific recognition

Box 7.1

Block copolymers

Block copolymers contain different monomers in each chain, but these are not randomly distributed but form blocks. Fig. 7.2 and Fig. 7.3 are concerned with linear di-block copolymers. The synthesis of such blocks requires controlled polymerization.

If the two monomers are incompatible, microphase separation occurs. The degree of incompatibility is proportional to the chain length. The parameter χN (with χ the Flory-Huggins interaction parameter and N the chain length) quantifies incompatibility.

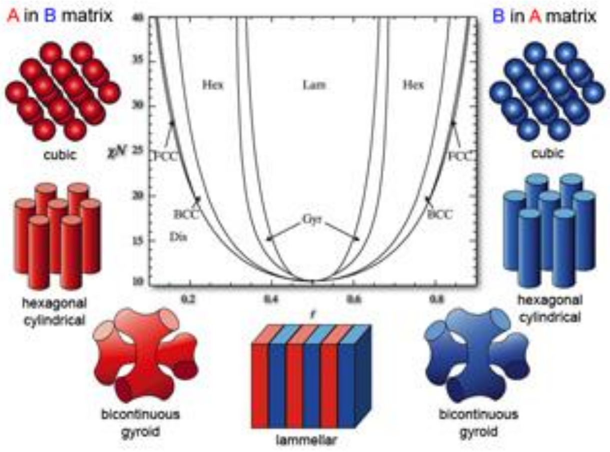


Fig. 7.2

For linear block copolymers, the phase diagram is known to be a function of block length ratio and incompatibility.

www.physics.nyu.edu/pine/research/nanocopoly.html.

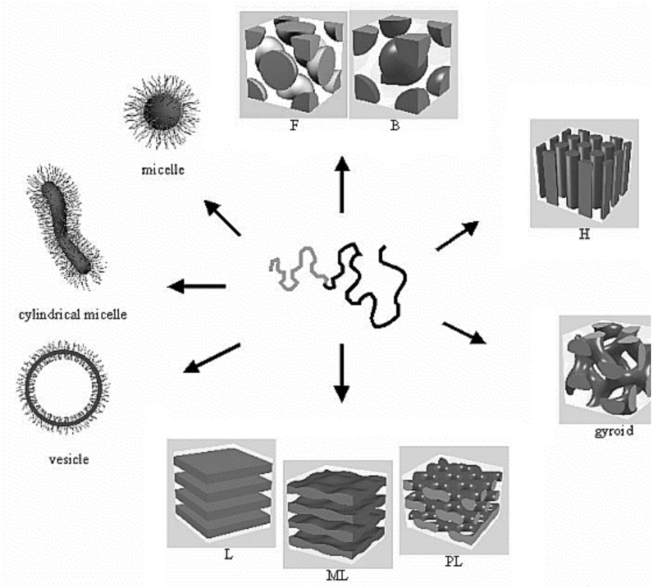


Fig. 7.3

Depending on block length ratio, incompatibility, degree of branching, and other properties, block copolymers form different morphologies.

www.chemie.uni-hamburg.de/pc/sfoerster/bilder/forschung/phasen.gif

We continue with comments on the different mechanisms

a) Frustrated incompatibility

When large molecules (such as polymers or surfactants) have domains that would segregate from each other but cannot because of a tight bond (usually covalent), the domains in question form at the microscopic level. Examples of such systems, in addition to surfactants, are block copolymers. Block copolymers are larger and more inert than typical surfactants. The self-assembled structures are often more stable than structures made from comparable surfactants.

Amphiphilic block copolymers play a certain role in controlled drug delivery. The polymersomes form nanocapsules that transport the active ingredient. They are referred to as "polymersomes" in reference to the structurally related liposomes. Polymersomes are often slightly larger than liposomes. The viscosity of the lamella is higher than the viscosity of the lipid lamella.

In the bulk, block copolymers produce what is known as microphase separation. This refers to "superstructures" from the different phases. **Fehler! Verweisquelle konnte nicht gefunden werden.** discusses microphase separation.

b) Topological constraints

Topology is a branch of mathematics that deals with the connectivity of objects. An often cited example is the difference between a torus and a pretzel. These two objects cannot be deformed into each other without cutting either the torus or the pretzel somewhere. A ring is topologically different from a thread with two ends. A membrane with edges is topologically different from a spherically closed membrane.

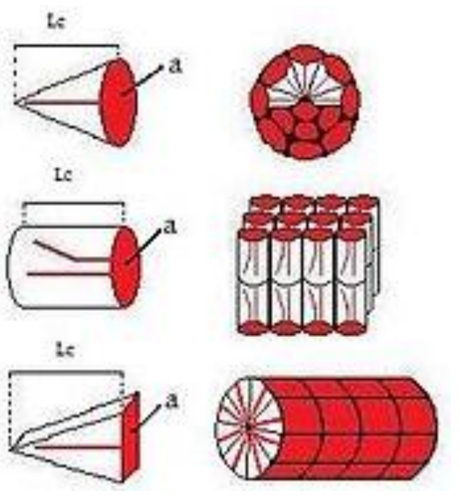


Fig. 7.4
The "shape parameter" is given as aL_c/V_c with a the area of the head group, L_c the length of the molecule, and V_c its volume. The larger the shape parameter, the larger the spontaneous curvature.
en.wikipedia.org/wiki/Thermodynamics_of_micellization

Lipid membranes rarely have edges because alkyl chains would come into contact with water at the edges. They are closed in one way or another. They often form spherical vesicles and thus divide the space into an inside and an outside.

When two vesicles fuse, this changes their topology. The process of "vesicle fusion" is carefully controlled by nature.

c) Spontaneous curvature

A spontaneous curvature of a lamella has similar consequences as the prohibition of an edge. The spontaneous curvature leads to spherical or cylindrical structures. The shape parameter of surfactants (Fig. 7.4) quantifies their tendency to take up a packing with curvature.

In the case of vesicles, both spontaneous curvature and the surface-to-volume ratio come into play (Fig. 7.5). The surface of a vesicle is largely solid. The volume is mainly fixed by the salt content (by the osmotic pressure).

d) Helicity

Helicity is self-limiting in two different ways. First, there can be attractive interactions inwardly, but not outwardly. This is the case with DNA. A second effect is somewhat more subtle. The inner fibers of a rope are barely bent ("bend") but twisted ("twist"). The reverse is true for the outer fibers. Nano-fibrils with a strong tendency to twist will form slender, twisted fibers instead of attaching themselves to an existing, thick fiber. This mechanism plays some role in the cellulose fibers in trees. Cellulose in trees takes the form of numerous slender fibers (“microfibers”) rather than thick rods

Due to their shape, cylindrical objects do not easily crystallize. Planar objects are packed into regular arrays more easily. This is why amylose (a component of starch) is readily soluble, whereas cellulose is not. The two polymers differ only at a single, seemingly insignificant site (the α- and β-glycosidic bond, respectively). The consequences of this small difference for packing are far-reaching.

e) Short-range attraction combined with long-range repulsion.

If an atom (more generally, a particle) interacts attractively with its neighbors, only (call number of neighboring particles z), but repulsively with all members of an aggregate (Fig. 7.7), the total energy of the aggregate E_{tot} can be written as:

$$E_{tot} = N\left(-Az + BN\right) = -AzN + BN^2 \qquad \text{Eq. 7.1.1}$$

N is the number of particles in the aggregate. *A* and *B* quantify the attractive and the repulsive interaction. In Eq. 7.1.1, the first term will dominate for small *N*, while the second term will dominate for large *N*.

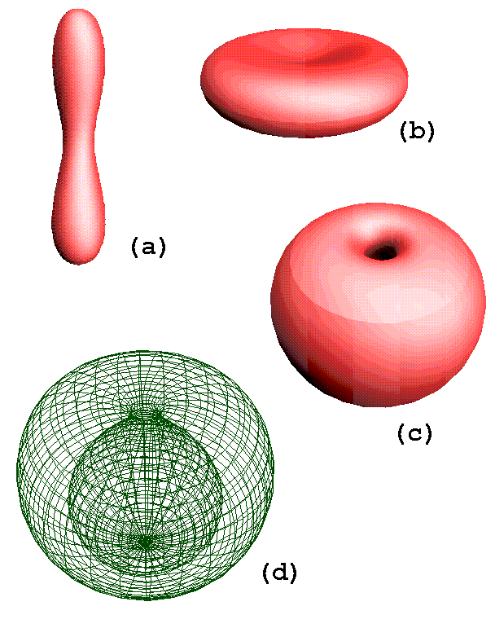


Fig. 7.5
For vesicles, the interplay between the surface-to-volume ratio and the membrane elasticity governs the shape.
www.maths.leeds.ac.uk/Applied/research.dir/Bio/biological.html

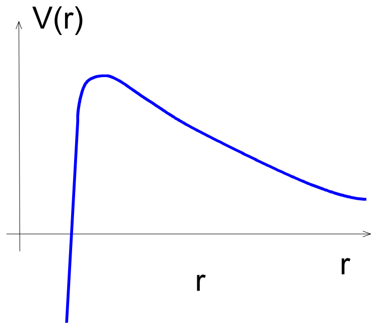


Fig. 7.7
Short-range attraction and long-range repulsion favor the formation of spatially confined aggregates.

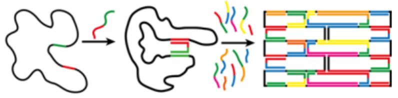


Fig. 7.6
On DNA origami.
www.spektrum.de/lexikon/biologie-kompakt/transfer-rna/11975

The long-range interaction in nature is often the electrostatic repulsion. Short-range in this context is the hydrophobic interaction and hydrogen bonding. This mechanism helps in the packing of globular proteins. Proteins are held together internally by the hydrophobic interaction and H-bonds, among other mechanisms. However, the charge on the proteins prevents the neighboring protein molecules from attaching to each other and precipitating. (For ordinary polymers, it is quite exceptionally difficult to obtain collapsed globules consisting of single chains). The same mechanism also causes *atomic nuclei* to contain no more than about 100 protons. The short-range interaction here is the nuclear force, the long-range interaction is the electric repulsion between protons (Fig. 7.7). Thus, the principles of self-organization are also at work in nuclear physics.

f) Pairs of molecules that recognize each other.

Self-organization has links to supramolecular chemistry, but – if defined as above – is not the same. Specific recognition occurs between *two* molecules (such as an antibody and an antigen). These two molecules (an example being an enzyme-substrate complex, ES) are not usually called a self-organized structure. If the specific recognition occurs between two chains, these may form a helix and those helices actually do count as “self-organized” (see the discussed under d).

A few side remarks:

- The examples above referred to aggregates in the bulk phase. "Self-assembled monolayers" (SAMs) usually refers to layers of thiol-terminated alkanes on gold. Here, a substrate stabilizes a structure that nevertheless exhibits self-organization (Fig. 5.12).
- In the above cases, equilibrium is always reached. This state is must be found and if the self-organized structure is complicated, finding it may be problem. This problem exists in protein folding. Base pairing of DNA also succeeds only with sufficiently small strands (oligonucleotides) or single nucleotides are attached to single-stranded DNA one-by-one. The DNA origami process (Fig. 7.6) illustrates this. Here, a self-assembled system is formed by base pairing, but the linking elements ("staples") are oligonucleotides. The staples do find their targets because they are small.
- Molecules might be viewed finite-size aggregates of atoms (similar to atomic nuclei being aggregates of protons and neutrons). Key to finite size are – first – the covalent bond and – second – the presence of monovalent atoms (often H, also F, Cl, and Br).
- Aggregates of finite size are needed to achieve complexity (linked to emergent behavior and life).
- Colloidal spheres have been turned into well-defined *super*structures by coating the surface with single-stranded DNA, which matched complementary DNA strands on other spheres. That works, but is complicated. Also, those superclusters are too large to achieve thermodynamic equilibrium. And: There was no strong case for application.

Remember

- Self-organization denotes the spontaneous formation of finite aggregates at equilibrium. Self-organization may lead to systems capable of emergent behavior.
- Examples of self-assembled systems include DNA, micelles, vesicles, the biomembrane, and self-assembled monolayers (SAMs) of alkane thiols on gold.
- Principles of self-organization include frustrated incompatibility, helicity, short-range attraction combined with long-range repulsion, topological constraints, and spontaneous curvature of membranes.
- Limited aggregates in this sense include the atomic nuclei and the molecules.
- Finite-size aggregates can also form by kinetically limited growth. Many nanoparticles and nanorods are produced in that way. These are not self-organized systems in the narrow sense.

7.2 Side remark: Langmuir layers

We return briefly to Fig. 5.4. On the right-hand side, the surface tension was largely constant because micelles form in the subphase at high surfactant concentration. This argument assumes that there is indeed a transport equilibrium between the surfactant layer at the surface, the dissolved surfactants, and the micelles. For long-chain surfactants, this is often not the case because they are not sufficiently soluble. Such molecules can be "spread" on the surface. In this way, the amount of substance can be fixed. One usually dissolves the molecules in an organic solvent (e.g. chloroform) and applies a drop of the solution to the surface carefully enough to not let the drop dive to the bottom. The drop spreads quickly, the solvent evaporates, and the surfactant layer (the "Langmuir layer") remains.

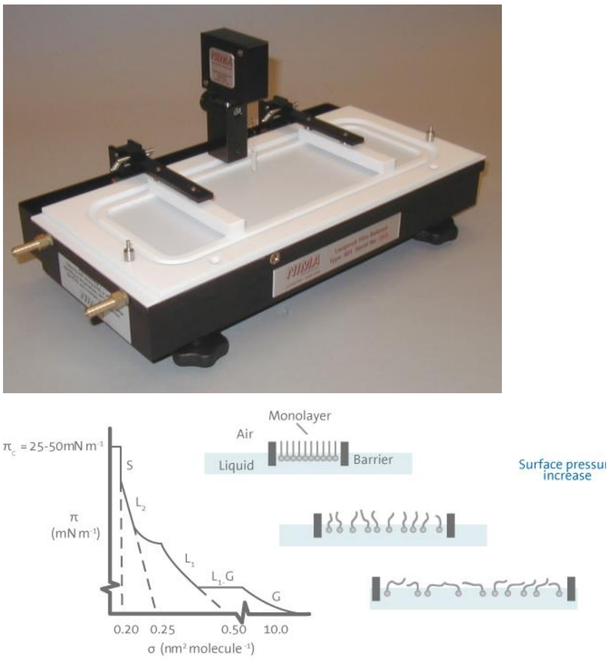


Fig. 7.8
The Langmuir trough and the πA diagram
Top:
www.engineerdir.com/product/catalog/16626/index1.html
Bottom:
www.biolinscientific.com/ksvnima/technologies/

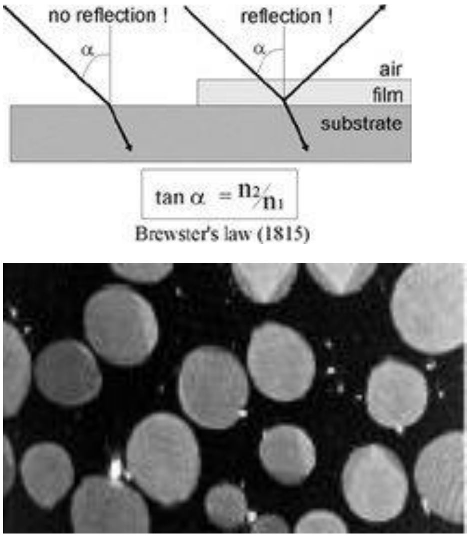


Fig. 7.9
The coexistence of two phases can be easily detected with the Brewster angle microscope. At the Brewster angle, the reflectivity of the surface is almost zero. A molecularly thin layer changes the reflectivity only minimally, but because this change occurs against a black background, it is clearly visible. In the case shown below, one phase forms round inclusions in the other phase.
Above: www.accurion.com/brewster-angle-microscopy
Below: users.otenet.gr/~garof/Bam/

In this way, one can create a two-dimensional systems and investigate their thermodynamics. The "concentration" is varied simply by changing the surface are with a movable barrier. One dips a sheet of paper (a Wilhelmy plate, Fig. 2.1) into the liquid and measures the surface tension. This experimental arrangement is called “Langmuir trough” (Fig. 7.8). In this context, the difference between the surface tension and the surface tension of pure water is viewed as as a 2D pressure and is called π . The πA diagram (bottom in Fig. 7.8) is an analogto the pV diagram in 3D.

Langmuir films are of interest for basic research because they allow to investigate the differences between 2D and 3D systems with regard to their phase behavior (Fig. 10.7). Furthermore, Langmuir films can be transferred to solid substrates by slowly dipping the substrates into the water phase and pulling them back out. Multilayers may formed, also called “Langmuir-Blodgett layers”. The physics of Langmuir films is of much relevance for wet foams (Section 8). Foam lamellae can be understood as an arrangement of two Langmuir layers with an intermediate water phase. Finally, one wonders whether the peculiarities revealed by these studies are relevant to the physics of the biomembrane. Critics deny this and call this line of research “biomimetic”, to be distinguished from “biological”.

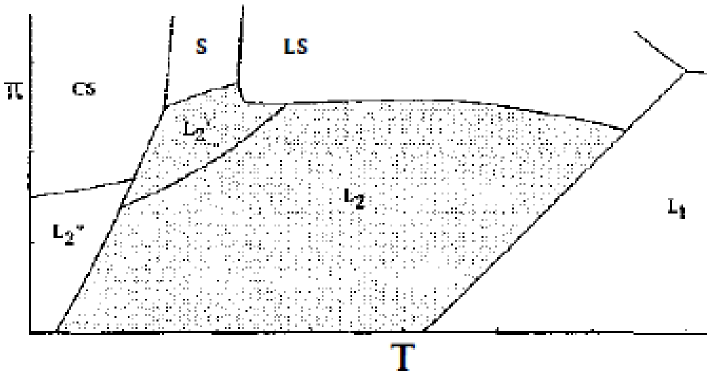


Fig. 7.10
Generalized monolayer phase diagram for long-chain acids, acetates and ethyl esters.
Bibo, A.M. et al. JPC 95, 5591 (1991).
The phases differ in lateral packing, crystallinity of the chain, and inclination angle of the chains. The characterization of the phases is done by X-ray diffraction (on molecular monolayers).

A few further comments on Langmuir films:

- If the πA diagram shows a horizontal line, one suspects phase coexistence (compare to the horizontal line in the pV diagram of the van-der-Waals gas). This phase coexistence can be visualized with the Brewster angle microscope (Fig. 7.9). One can also visualize it by staining with dyes. The horizontal lines in the πA diagrams are often not strictly horizontal because the dispersed domains have a vertical dipole and repel each other.
- The region at high area and low pressure would be thought to be a gas phase. However, a gas phase is very difficult to observe because of the strong interactions between the alkyl chains. It is an "expanded" liquid phase.
- One can visualize the nature of the different phases with X-ray diffraction. To do this, one must use synchrotron radiation under grazing incidence, because the sample is quite extraordinarily thin.
- X-ray diffraction on Langmuir films reveals a wide variety of phases (Fig. 7.10). There is a deeper reason for this. In 3D, a favorable packing has to be achieved with about 12 neighbors in all three dimensions, in 2D only with about 6 neighbors in two dimensions. The latter is easier because "frustration" occurs less often. Therefore, there is a greater variety of phases.
- The different phases can form complex superstructures of all kinds. These are aesthetically pleasing, but the technical relevance is not always obvious.
- In addition to structure, one can also study dynamics on Langmuir films. "Dynamic measurements" here means 2D rheology. 2D rheology is important for foams. But it could also mean – at least in principle – dielectric spectroscopy. (NMR is impossible, because the sensitivity of NMR is not sufficient for monolayers).

Remember

- Water-insoluble surfactants form two-dimensional systems at the water/air interface, the “Langmuir films”. These can be transferred to solid substrates (Langmuir-Blodgett films).

- The πA diagram is the 2D analog of the pV diagram. The horizontal areas indicate phase coexistence.
- The phases can be visualized with Brewster-angle-microscope.
- X-ray diffraction allows to characterize the phases. There are more different phases than in 3D.
- A similar diversity of phases is suspected for the biomembrane (keyword: "rafts").

8 Bubbles and foams

A distinction must be made between aqueous foams and foams with a solid (possibly also soft, but elastic) continuous phase. In foams with an elastic matrix, the outstanding feature is the low density. In semiconductor electronics, the low dielectric constant is an additional feature. In polymer foams, the volume compressibility (dominated by the high compressibility of the gas phase) is added. Foams are also used for thermal insulation.

First, some remarks on isolated bubbles (to be distinguished from air spaces bounded by lamellae, as in foams):

- Bubbles are subject to Ostwald ripening. Nanobubbles are unstable. Reminder: Ostwald ripening is prevented in miniemulsions by adding a "hydrophobe". There is no such mechanism for bubbles.
- Bubbles are subject to *cavitation*. This is a collapse driven by Laplace pressure. It leads to extremely high local temperatures and pressures. In a technical context (e.g. ships), cavitation must be avoided because it easily creates damage. Ultrasound can cause cavitation. Some effects of ultrasound (such as the disaggregation of clumps of particles) are linked to cavitation.
- Nanobubbles may be stabilized by certain surfactants. Such nanobubbles are contrast agents for ultrasonic imaging due to their high compressibility. Air pockets are sometimes advantageous for adhesives and sealants (e.g. on window panes) due to their compressibility.
- Adsorption of bubbles onto mineral particles is significant in the context of *flotation.* Special surfactants are used, which promote the selective adsorption of bubbles onto certain particles and thus enable the separation of powders according to chemical constitution.
- The van-der-Waals interaction between bubbles is attractive. The effective Hamaker constant is positive.
- The surfaces of bubbles are often charged due to preferential adsorption of ions. Therefore, bubbles usually repel each other (despite the attractive vdW interaction).
- The van-der-Waals interaction between bubbles and solid surfaces often is repulsive. This occurs when the Hamaker constant of the liquid is lower than the Hamaker constant of the solid. One then calculates an "effective Hamaker constant"[49] for the interaction between bubble and solid, which is negative. (This is a problem for flotation.)
- When stabilization of bubbles is achieved with lipids, the high viscosity of the lipids is important in addition to the lowered surface tension.

Among *aqueous foams*, an economically important example (and well-studied example) is foam on beer. Foam is also found in washing processes, e.g. in car washes. Another application example is fire protection. Otherwise, foam formation in working fluids is rather undesirable. Foams consume space. There are additives for foaming as well as for defoaming, the second application being basically the more important.

---

[49] Write $A_{123} = (A_{11}^{1/2} - A_{22}^{1/2})(A_{33}^{1/2} - A_{22}^{1/2})$ with "1" the solid, "2" the liquid, and "3" the air.

Other comments on foams include:

- The soap lamellae separating the air spaces usually contain a water layer in the center. The thickness of this layer is variable. It can, for example, be of the order of the Debye length. It can also be even thicker if polyelectrolytes are added. The physics of the soap lamellae is very different from the physics of the biomembrane.
- Foams are destabilized via drying, Ostwald ripening, and coalescence.
- To stabilize foams, one stabilizes the lamellae. Polymers, large molecules (e.g. proteins) and also particles are often used. This is where 2D rheology comes into play.
- Foams are destabilized with *defoamers*. *Defoamers* are usually amphiphilic substances that form lamellae poorly (i.e., do not tend to self-organize). These initially adsorb to the water-air interface, but then tend to form droplets themselves. On the way, the lamella collapses.

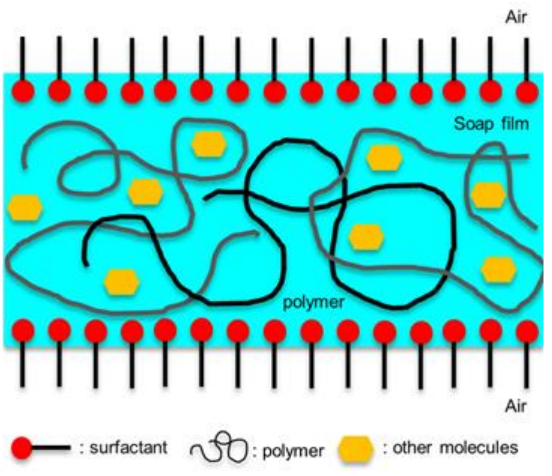


Fig. 8.1
On the structure of soap films. If there are repulsive interactions between the two surfactant layers, this lowers the negative energy of the lamella. Then the capillary pressure of the soap bubble is slightly smaller than $4\gamma/r$.
pubs.sciepub.com/wjce/4/2/2/

Remember

- A distinction is made between wet and dry foams (the latter often made of polymers).
- Foams have low density, high compressibility, and low dielectric constant.
- Stabilizing nanobubbles is difficult because of the high Laplace pressure.
- Wet foams are compartmentalized by lamellae. The boundaries between the lamellae are the plateau borders. The lamellae are a few nanometers thick. In the water layer, forces between charges play a major role.
- Foams are destabilized by drainage, coarsening and coalescence.
- Defoamers attach to the lamella, but are themselves unstable (do not tend to self-organize).

9 Remarks on tribology

Tribology is about *friction* and *wear*. Friction can be desirable (adhesion) or not (reduction of energy consumption). Wear may be desirable (machining) or not. Where wear is unavoidable, it should be gradual and controlled (tires). Sometimes, in addition to friction and wear, noise generation also plays a role (whispering asphalt, freight car brakes).

Tribology is a cross-cutting issue across engineering with huge financial implications (almost like corrosion). Tribology is addressed in an interdisciplinary manner. Phenomena from many different spatial and temporal scales contribute. There is a wide variety of material combinations and problems in detail.

Examples of tribosystems are:

- Rolling bearings
  e.g. ball bearings or bearings with rollers
  Rolling bearings are favorable in frictional resistance. However, there are high stress concentrations at the contact points. The modeling is done in the framework of the *lubricant film theory*. Wind turbines are currently operated with rolling bearings.
- Plain bearing lubricated
  e.g. piston in engine
- Plain bearings without lubrication (Teflon, PEEK)
  In this case, a "transfer film" is formed from the abrasion of polymers, which facilitates sliding.
- Wheel-rail contact, tire-road contact
  Contacts under rolling friction, non-lubricated.
- Read head above the hard disk
  Here, the air takes over the role of the lubricant. Intermittent contacts occur. For this reason, the hard disk is coated with a thin layer of perfluoroethers (e.g. "ZDOL").
- Biolubrication
  Here, there is never concentricity; the relative motion always comes to a halt.
  The tribo-surfaces are formed by the extracellular matrix. Hyaluranic acid plays a major role. Hyaluranic acid may be thought of as a polymer brush (a polyelectrolyte brush). Such brushes are locally fluid structures, but they are not washed away due to terminal attachment to the substrate. The interaction *between* brushes is always repulsive for entropic reasons. This is favorable for friction.
- MEMS devices
  MEMS stands for microelectomechanical systems. MEMS devices are often made of silicon. One uses the "CMOS technology", which is well known from the semiconductor industry. The motion is often induced and read out piezoelectrically. Examples of MEMS are the AFM tip, acceleration sensors for airbags, and systems of small mirrors for projection systems.

  In principle, MEMS devices can also be used to construct rotor-shaft systems (micromotors). However, these have so far had a problem with tribology. A related problem sometimes occurs with irreversible adhesion (e.g. promoted by capillary forces).

This concludes the list of examples. We turn to the basics. *One must fundamentally distinguish between friction in liquids and friction between solids.*

9.1 *Friction in liquids*

In liquids, the frictional force is proportional to the velocity if the velocity is not too high.[50] The force-velocity relation is *linear*. We treat three geometries: the sphere in the fluid, the sphere in the fluid near a surface, and the flow in a pipe.

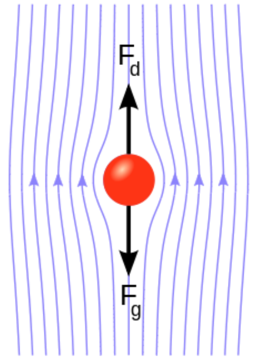


Fig. 9.1
To the friction according to Stokes
file.scirp.org/Html/4-

The friction law for spheres in Newtonian fluids is Stokes' law (Fig. 9.1):

Eq. 9.1.1

$$F_{fric} = 6\pi r \eta \mathrm{v}$$

The friction force is proportional to the radius, that is to the size, as opposed to the surface area. This result is general result. For non-spherical particles, one writes $F_{fric} = 6\pi\eta r_h \mathrm{v}$ and calls the length defined in such a way the *hydrodynamic radius*.

The friction force can be compared with other forces. A typical such force would be gravity, given as

Eq. 9.1.2

$$F_{grav} = \left(\rho_p - \rho_{liq}\right) g \left(\frac{4\pi}{3} r^3\right)$$

Putting gravity and the frictional force in equilibrium, we calculate a sedimentation velocity, given as

Eq. 9.1.3

$$\mathrm{v} = \frac{2r^2 g(\rho_p - \rho_{liq})}{9\eta}$$

Thus, the sedimentation velocity becomes small on small scales. More generally, the influence of friction often becomes large on small scales.

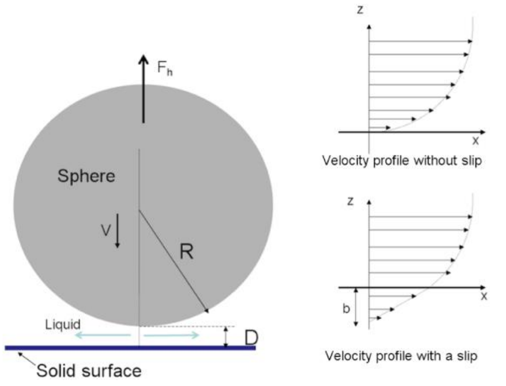


Fig. 9.2
When a sphere approaches a planar surface, the frictional force increases inversely proportional to the distance. This may be different in the presence of slip.
(*R* in this figure is called *r in* the text).
iopscience.iop.org/article/10.1088/0953-8984/25/18/184003 Maali, A.; Bhushan, B., *J. Physics-Condensed Matter* **2013,** 25, (18).

When a sphere approaches a surface, the frictional force increases in proportion to the inverse distance, *D*, in the limit of $D \ll r$ (Fig. 9.2). The Reynolds law state that

Eq. 9.1.4

$$F_{fric} \approx 6\pi\eta r \mathrm{v} \, \frac{r}{D}$$

Since the frictional force tends towards infinity shortly before the impact, the impact never occurs – within this model. This is put to use in ball bearings.

For the flow in a pipe with radius *r* and length *L* the Hagen-Poiseuille law applies:

---

[50] At very high speeds, it is no longer the viscosity that is constant, but the stress. This is referred to as "limiting shear stress". The conditions are then similar to the friction of solids on solids according to Amonton's law. The viscosity then scales like v^{-1} .

Eq. 9.1.5

$$\mathrm{v} = \frac{r^2}{8\eta L}\Delta p$$

Δp is the pressure difference. The smaller the radius, the larger the influence of friction. A side remark: Capillary forces also increase on small scales. If there is an equilibrium of forces between friction force and capillary force, it is not clear a priori which one is stronger on small scales.

For all these geometries, a *"no-slip" condition was* postulated at the surface. This is often appropriate, but not always. For gas bubbles, it is different. For gas bubbles, the factor of 6π in Stokes' law is replaced by a factor of 4π. Slip is even more significant in the case of flow through a capillary than in the Stokes' law. In the presence of slip, there is a "slip length" (usually called *b*, cf. Fig. 9.2). One can define an effective capillary diameter as $r + b$. Slip sometimes results on hydrophobic surfaces from "nanobubbles" and "nanopancakes" of air. Some small animals lower their hydrodynamic resistance in water by thin layers of air on their surfaces.

The slip length can be several 10 nm, especially for water on hydrophobic surfaces. The roughness of the surface, the density of the water near the surface (which has to do with the interactions with the wall) and whether there are ice-like structures at the surface are important. One naturally thinks of an engineering application in terms of reducing hydrodynamic drag in fine capillaries. However, consider this: this large slip length occurs on hydrophobic surfaces. If at the same time capillary force is supposed to drive transport, you have a trade-off. In porous media, slip lowers friction, but you still have the viscous dissipation that results from the complex structure. This issue can be illustrated by looking at the drag of an air bubble. (At the surface of the bubble, the slip length is infinite.) This is given as $F/\mathrm{v} = 4\pi\eta r$, to be compared with the value $F/\mathrm{v} = 6\pi\eta r$, which results according to Stokes' law for a sphere with a no-slip condition. "slip" or "no-slip" accounts for only a factor of 2/3 here. This is different for the flow in nanotubes. Very low flow resistances were found for the flow of water through carbon nanotubes.

Do not confuse this form of slip with the sliding of a polymer melt against the wall in an extruder. The latter is a consequence of shear thinning. Viscosity can drop in polymers when the shear rate is high. This happens mainly at the edge of the tube. Then a positive feedback occurs (η decreases, shear rate increases, η decreases further, shear rate increases further ....) until plug flow occurs. We know plug flow from toothpaste.

Further, do not confuse this form of slip with what the sharkskin does. The sharkskin is rough and thus lowers the diameter of the rolls, which result from turbulence. (Such structures are also used in airplanes.) Turbulence was left out here. (Turbulence takes place on large scales).

A side remark: Elastohydrodynamic friction (EHD friction) is friction in fluids. Nevertheless, the friction force is proportional to the normal force and largely independent of the velocity (see below). "Lubricaton theory" applies. Of some importance in this context is the fact that the viscosity of normal oils increases under

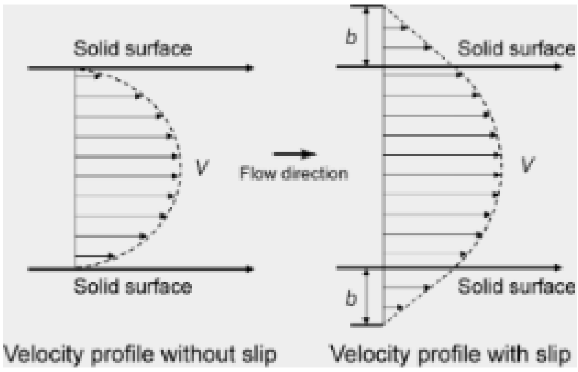


Fig. 9.3
About the slip and the slip length
www.beilstein-journals.org/bjnano/single/articleFullText.htm?publicId=2190-4286-2-9

pressure. High pressure occurs locally before the surfaces come into contact. The high viscosity causes dynamic stabilization of the fluid film and thus reduces wear. Water is a poor lubricant because this viscosity increase with increasing pressure is virtually non-existent in water.

Of further importance for the lubricating film theory is that at high shear rates the viscosity of most fluids decreases. High shear rates are present in EHD friction in the gap. In the limiting case, the viscosity is inversely proportional to the shear rate. The shear stress (viscosity × shear rate) then becomes independent of the shear rate. As soon as this occurs, it makes sense to speak of a coefficient of friction for lubricated friction as well.

9.2 Friction between solid surfaces

Friction between the surfaces of solids is often described by a remarkably simple law whose justification is anything but trivial. It is Amonton's law[51] :

Eq. 9.2.1

$$F_{\|} = \mu\, F_{\perp}$$

$F_{\|}$ and $F_{\perp}$ are the tangential and normal force. μ (a number of the order of 1) is the *coefficient of friction*. Here, the "static" coefficient of friction (at the onset of sliding) is slightly higher than the dynamic coefficient of friction ($\mu_S > \mu_D$). Remarkably, the friction force is independent of the velocity and the apparent contact area in the Amontons scenario. The behavior is nonlinear. People got used to these laws. Some surprise would be appropriate. The Amontons laws are only are approximate. More precise measurements reveal a small velocity dependence of the friction force.

The Amontons laws were explained by Bowden and Tabor with roughness. There are "load-bearing asperities". Their area is just large enough that the local normal stress corresponds to the yield stress (σ_y). The true contact area is therefore proportional to the normal stress. During sliding, these local contacts constantly break. They do so at a critical tangential stress, σ_c. Bowden and Tabor write $\sigma_c = \mu\sigma_y$ and arrive at

Eq. 9.2.2

$$F_{\|} = A\sigma_c = A\mu\,\sigma_{\perp} = A\mu\,\frac{F_{\perp}}{A} = \mu\, F_{\perp}$$

Note the relationship between friction and fracture mechanics. Dry friction is basically more closely coupled to wear than lubricated friction.

When the friction coefficient becomes very small, "superlubricity" is present. Superlubricity requires atomically smooth incommensurable crystal surfaces.

A coefficient of friction also exists for lubricated friction, i.e. friction in the presence of fluids. The Stribeck curve (Fig. 9.4) shows the coefficient of friction for lubricated friction as a function of a (suitably normalized) sliding velocity. At low velocities, Amonton's law generally applies. At high speeds, a lubricating film keeps the two surfaces at a distance. The tangential force is proportional to the normal force because the gap width increases with speed. This feedback mechanism makes the

---

[51] Amontons (with s) is the proper noun. The s is not a genitive s.

system non-linear. The region in question goes by the name elastohydrodynamic friction (EHD friction). The main problem in EHD friction is wear in intermittent contact.

The area of "mixed lubrication" is particularly problematic for wear. The service life of car engines therefore depends to a large extent on the number of starts.

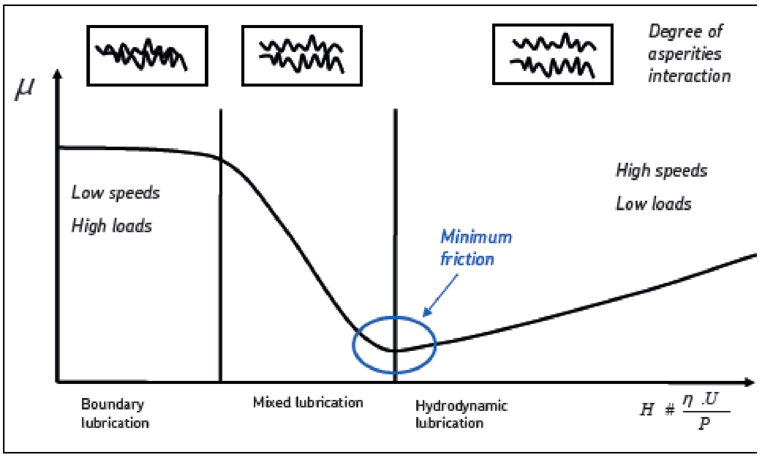


Fig. 9.4
The Stribeck curve
www.ibtinc.com/grease-analysis-in-the-field-1/

9.3 Comments on practical tribo systems and lubricants

The following are some remarks on practical tribology:

- Hard surfaces are of great importance for lubricated surfaces (ball bearings). The surfaces are often specially hardened (e.g. by nitriding).
- In engines, there is "running-in" in the sense that carbon from the lubricating oil is deposited in the areas of the steel near the surface.
- If there is abrasion, it is important that it is controlled and gradual.
- Zinc dialkyl dithiophosphate (ZDDP) is often added to lubricating oils. This creates "anti-wear layers" at the areas in danger of damage.
- Polymers, that swell at high temperature, are often added to lubricating oils, causing the temperature-viscosity profile to be flat.

<u>Remember</u>

- Tribology is a cross-disciplinary topic.
- Tribology is complex because many space and time scales are involved.
- Frictional forces become large on small scales.
- Friction in fluids is linear in velocity at sufficiently small velocities ($F \propto$ v, Stokes' law applies).
- Thin layers of air can lead to slip at solid surfaces in liquids.
- Friction on solid surfaces is nonlinear ($F_{\parallel} \approx \mu F_{\perp}$).

10 Colloids

The word "colloidal" originates from Greek and can be translated as "glue-like". Colloids sometimes are sticky. At the time, the distinction between colloids and polymers had not been well established.[52] Polymers are a bit more sticky than colloids. The stickiness of a material is closely related to its flow properties (to its rheology). Colloids have interesting flow properties. Colloids (and also polymers) are "complex fluids." The field is sometimes called "soft matter".

According to the IUPAC definition, "colloidal" denotes an "object, which at least in one of the dimensions lies between 1 nm and 1 µm". This definition is broad and not very meaningful with respect to the underlying phenomena. In the following, we restrict our interest to small dispersed solid objects in a liquid environment. These, of course, constitute a subset of all colloids. We do not discuss foams, aerosols, lamellar structures, and solid/solid colloids.

With this restriction a more speaking definition is possible. Colloidal objects in this sense are larger than molecules, but still so small that the thermal motion (also: “Brownian motion”) brings about a *thermodynamic equilibrium.* This definition also leads to a maximum size in the order of 1 µm (because larger particles sediment and have slow dynamics).

The thermodynamic equilibrium might not reached

- on short time scales (for instance in fast shear flow)
- in the presence of Ostwald ripening (only with liquid/liquid emulsions). Ostwald ripening can be prevented by using a "hydrophobe".
- when sedimentation beats diffusion
- when aggregation is not prevented

Thermodynamic equilibrium occurs on the time scale of diffusion, the latter given as

Eq. 9.3.1

$$\tau_{diff} = \frac{L^2}{D} \approx \frac{a^2}{k_B T / (6\pi\eta a)}$$

D is the diffusivity. In the second step, the Stokes-Einstein relation was used. Furthermore, the particle radius, a, was used for the characteristic length, L.[53] (One might also have used four particle radii). Inserting (e.g.) numerical values of $k_BT = 4\cdot10^{-21}$ J, $\eta = 10^{-3}$ Pa s, and $a = 100$ nm leads to $\tau_{diff} \approx 5\cdot10^{-3}$ s. This time scale can be compared, for example, with a shear rate, $\dot{\gamma}$. The shear rate has the dimension 1/s. The product $\dot{\gamma}\,\tau_{diff}$ also carries the name "Deborah number" De (a characteristic number from rheology, into which other characteristic times are sometimes inserted). For $De > 1$, a non-trivial rheological behavior can be expected. One such non-trivial behavior (among others) is shear thickening. If a dispersion has a high solids content, this dispersion will only flow slowly. If one tries to shear it quickly, "jamming" occurs. There is transient solidification.

Regarding stability against sedimentation: the interplay between sedimentation and diffusion leads to a Boltzmann distribution of particles along the vertical coordinate z:

---

[52] Polymers contain linear chains, which is reflected in rubber elasticity and the sluggish flow caused by entanglements. Colloids do not exhibit rubber elasticity.

[53] In colloid science, the sphere radius is mostly called a (rather than r or R). We folow the common usage.

Eq. 9.3.2

$$c(z)=c_0\exp\left(-\frac{mgz}{k_BT}\right)=c_0\exp\left(-\frac{z}{\xi}\right)$$

ξ is the characteristic height. $m = \Delta\rho\ 4\pi/3\ a^3$ and g are the mass and the acceleration constant due to gravity. One finds

Eq. 9.3.3

$$\xi=\frac{k_BT}{\frac{4\pi}{3}a^3\Delta\rho g}$$

With $\xi = 1$ m and $\Delta\rho = 0.1$ g/cm^3 one arrives at a critical particle radius of about 10 nm. Why then the critical size of 1 µm?. This critical size arises from the density contrast often being is smaller and 0.1 g/cm^3 and, also, because there just about always is some thermal convection helping us to fight sedimentation.

As a side remark, we prove the Stokes-Einstein relation from the fact that the interplay of sedimentation and diffusion must lead to the Boltzmann distribution. The sedimentation flux J_{sed} is

Eq. 9.3.4

$$J_{sed}=c\mathrm{v}_{sed}=c\frac{F_{grav}}{\xi_{fric}}=c\frac{mg}{6\pi\eta a}$$

v_{sed} is the velocity, F_{grav} is the force of gravity. The concentration distribution $c(z)$ is subject to the Boltzmann law (Eq. 9.3.2). The diffusive flux is

Eq. 9.3.5

$$J_{diff}=-D\frac{\mathrm{d}c}{\mathrm{d}z}$$

We equate the two fluxes and use insert Boltzmann distribution for $c(z)$:

Eq. 9.3.6

$$J_{sed}=c\frac{mg}{6\pi\eta a}=J_{diff}=-D\frac{\mathrm{d}c(z)}{\mathrm{d}z}=-D\frac{\mathrm{d}}{\mathrm{d}z}\left(c_0\exp\left(-\frac{mgz}{k_BT}\right)\right)=D\frac{mg}{k_BT}c$$

Solving for D leads to $D = k_BT/(6\pi\eta a)$.

On aggregation: The vdW potential in the limiting case of $D_{surf} << a$ *is* given as $-Aa/(12D_{\mathrm{surf}})$ with A the Hamaker constant, a the particle radius and D_{surf} the distance between the surfaces. For fixed D_{surf} (e.g. 5 nm, corresponding to about two Debye lengths), the vdW potential increases with increasing a until one can no longer compensate it by any repulsive interaction.

Particles and aggregates of particles with sizes far above 1 µm carry the name "granular matter". Granular matter plays a major role in industry in the form of powders, granulates and bulk materials. For granular matter, one chooses a particle size larger 1 mm. One can then rely on gravitation and sedimentation (even in air, keyword: dust formation). Also, one can always bring about disaggregation by shaking. (Here inertial forces and adhesion forces are in competition. On large scales, the inertial forces win).

A side remark: Most colloidal spheres are too small for optical microscopy. People have developed special spheres, which

- are large enough to be resolved with optical microscopy
- have the same density as the liquid (→ no sedimentation)
- have the same Hamaker constant as the liquid (→ no aggregation)
- have the same refractive index as the liquid (→ no turbidity)

These spheres are made to be fluorescent, so that they can be imaged with a confocal microscope. The liquid is made viscous enough to slow down Brownian motion, so that it can be followed with the laser scanning confocal microscope (LSCM). This provides for a detailed picture of structure and dynamics. The vitrification and crystallization of colloids, for example, have been studied in this way.

Some numbers are interesting. Consider a total volume of 2 L. Let the solids content be 50% (so that the polymer phase has a volume of 1 L). Let the volume of a particle be about 100 nm×100 nm×100 nm. If the particle consists of polymer chains with 1000 segments each and if the volume of a segment is 0.1 nm^3, the sample contains 10^{18} particles, each particle contains 10^3 chains, and each chain contains 10^3 segments. Large numbers come into play three times. The numbers have different implications:

- The number 10^{18} may seem high, but it corresponds to only 1 µmol/L. (For comparison: water contains 50 mol/L.) The osmotic pressure of polymer dispersions is correspondingly small. We neglect the interactions (which is problematic at a solids content of 50%) and apply the van't Hoff equation ($\Pi = RTn/V$). Inserting the values (RT = 2.48 kJ/mol, n = 1 µmol, $V = 10^{-3}$ m^3) leads to a pressure of 2.4 Pa, corresponding to 0.024 mbar. For comparison: seawater (with a salt content of 3.5 wt%) generates an osmotic pressure of 30 bar.
- From the small number-density of particles it follows that latex blends (mixtures of different types of latex particles) have a strong tendency to segregation. If there are very weak favorable interactions between similar particles,[54] the - small - negative contribution of the mixing entropy to the free energy of the latex blend is overcompensated by the (positive) mixing enthalpy.
- The number 10^{18} shows that it is virtually impossible to produce significant amounts of material from particles that are created one by one as drops at the end of a capillary ("droplet-based microfluidics"). A nozzle can produce about 10^6 particles per second. Such quantities are at best interesting for medical applications.
- Since a particle contains 10^3 polymer chains, the interior of the particle may be considered as a normal bulk polymer phase. Anomalies that would result from the geometric constraints (from the "confinement") are not expected. The same conclusion is reached if one remembers that the surfaces between condensed phases are molecularly sharp. With a particle radius of 50 nm and a surface width of a few nm, most chain segments do not see the surface.

Remember

- According to IUPAC, colloidal systems have a size between a few nm (just above the molecular size) and about 1 µm. We restrict our interest to dispersed particles in a continuous phase, where

---

[54] The interactions between unequal particles then are less favorable than those between equal particles.

the particles are smaller than 1 µm. This upper limit results from the requirement that Brownian motion should prevent sedimentation. ("Diffusion beats sedimentation.")

- For much larger particles (> 1 mm), rapid sedimentation is certain and deagglomeration by shaking is always possible. These are the "granular media".
- In polymer dispersions, a sphere may well contain a large number of chains. One may call the interior a "bulk".
- Colloidal dispersions do not usually build up significant osmotic pressure because the number of particles – and the mixing entropy, as well – is small.

10.1 Examples of colloidal systems

- Fine powders can have colloidal properties. The manufacturing process is as "top-down" (to be distinguished from "bottom-up").[55] Grinding is expensive and rarely results in particles smaller than 50 nm due to reaggregation.

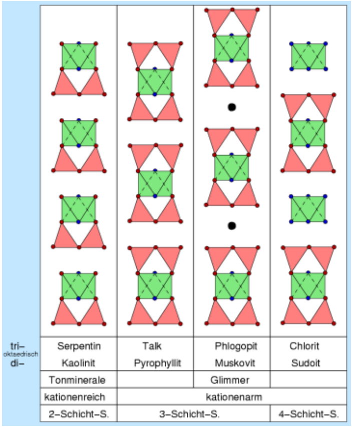


Fig. 10.1
Alumino-silicates often form layered structures.
ruby.chemie.uni-freiburg.de/lecture/silicate_7_5.html

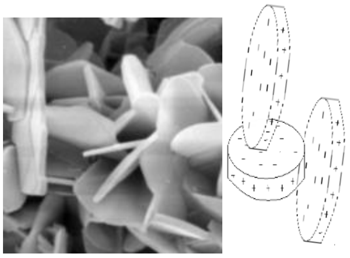


Fig. 10.2
Due to the different charge on edges and surfaces, clay minerals form the house-of-cards structure.
www.uni-kiel.de/anorg/lagaly/ECGA/Fluorhectorite.gif

Fig. 10.3
Micelles are often incorrectly depicted as spheres. There shape is more irregular.
www.uni-essen.de/~cua040/research.html

- Globular proteins (and antibodies[56], as well) are molecules, basically, but certain properties can be understood within the framework of colloidal physics. Proteins mostly are charged. There are free counterions. Unlike ordinary polymer dispersions, the charge varies on the surface and there are hydrophobic regions on the surface. Proteins are called "patchy colloids".

---

[55] Talking of "top-down": In semiconductor technology, photolithography is a top-down process. Attempts to support photolithography with self-organization processes have not yet been put into practice. In photolithography, structure widths of a few nm can be achieved.
Top-down structuring down to the nanometer range is possible with the "focused ion beam" (FIB). It is possible to cut parts out of a crystal (in vacuum) and to investigate these parts (mechanically, optically, magnetically) with respect to the influence of the size on the physical properties. This is a rich field, especially in mechanics, because the effects of defects (e.g., dislocation lines and their migration) are undermined.

[56] But not the membrane-bound proteins and the proteins that form fibrils.

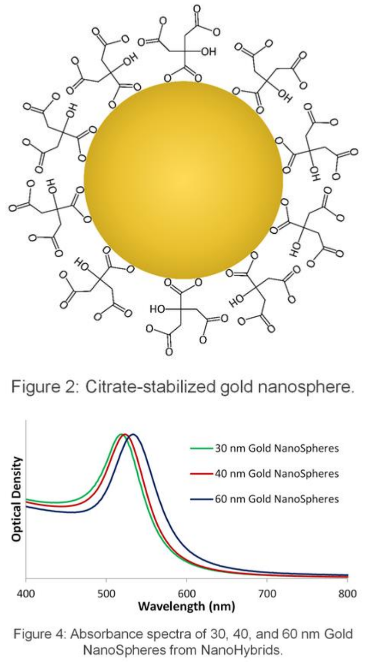


Figure 2: Citrate-stabilized gold nanosphere.

Figure 4: Absorbance spectra of 30, 40, and 60 nm Gold NanoSpheres from NanoHybrids.

Fig. 10.4
Top: Citrate-stabilized gold nanoparticles (AuNPs).
Bottom: Isolated AuNPs show the characteristic plasmon resonance in optical absorption.
nanohybrids.net/pages/gold-nanoparticles-for-lateral-flow-assays.

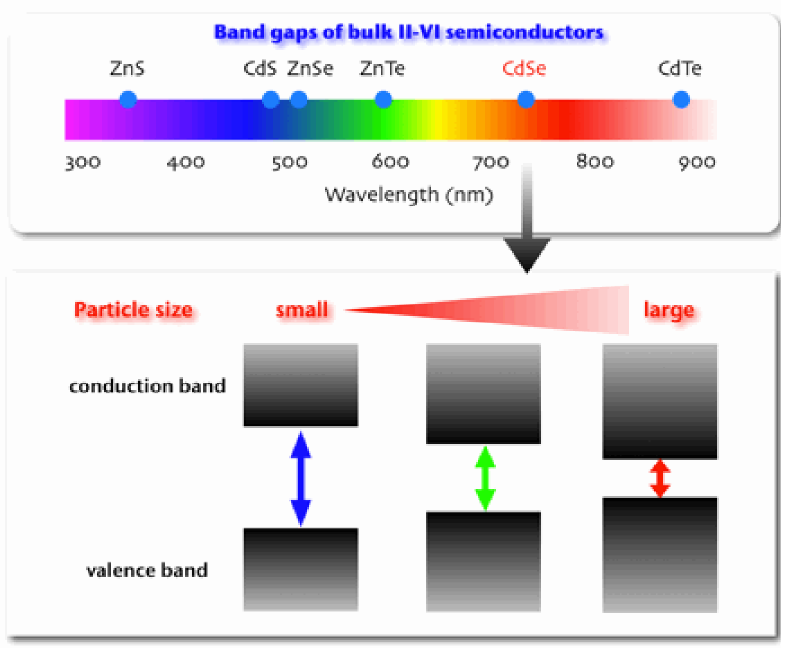


Fig. 10.5
In the case of semiconductor nanoparticles, the band gap depends on the diameter. (Remember the "particle in a box" from quantum mechanics.) Color can be tuned by size.
www.scl.kyoto-u.ac.jp/~teranisi/research/Opt-e.html

– Clay minerals ("layer silicates", "alumino-silicates", "nano-clays", Fig. 10.1) often form platelets. Examples are kaolinite and montmorrillonite. Mica is also a layered silicate, but usually is larger in diameter than a few 100 nm. The formation of these finite aggregates can be thought of as a process of self-organization (albeit far away from thermodynamic equilibrium). The layered silicates consist of corner-sharing tetrahedra (based on SiO_2) and octahedra (based on Al_2O_3). When a corner of a tetrahedron not shared, a net negative charge remains, which is compensated by cations. Why the tetrahedra and octahedra often form layers, is not obvious. (At times they form other structures, e.g., the zeolites.) The mechanism is robust in the sense that there are quite a few related structures and that small deviations in stoichiometry from ideal conditions are tolerated.
In the dry, the various layers are ionically bonded by the cations. In water, these physical bonds open up. The layers are "exfoliated". The size is in the range of a few nanometers. The edges are often positively charged, while the surfaces are negatively charged. If the charges are not screened, the house-of-cards structure is formed when the concentration in water is high enough (Fig. 10.2). The house-of-cards structure produces a weak gel. Clay minerals can be used as rheology modifiers. The charge at the edges is sometimes screened with phosphates. The gel then turns into a fluid.

Clay minerals bind cations on the surfaces. They are important for soil hydrology. Clay minerals are also important for ceramics.[57] Do not confuse clay minerals with cement, though.

- Polymer dispersions consist of polymer particles with a diameter of about 100 nm. They are used as binders, coatings, and adhesives. Emulsion polymerization is a mature production process.
- Association colloids consist of surfactants aggregated into superstructures. The most important example is the micelles (Fig. 10.3). However, lamellae and cylindrical structures also form. Closed lamellae form vesicles (if formed from lipids: "liposomes"). These vesicles are "nanocontainers".
- Gold nanoparticles (AuNPs, Fig. 10.4) can be prepared relatively easily from gold acid (see also point c in Sect. 7.1). The citrate creates a surface charge that stabilizes the particles against aggregation. A typical size is 10 nm. Gold nanoparticles - when isolated - have a bright red color due to the "plasmon resonance". Plasmon resonance is a collective excitation of quasi-free electron gas. Gold nanoparticles have already been used in church windows . Due to their color and easy surface functionalization, AuNPs are popular markers in bioanalytics.
- Semiconductor nanoparticles (Fig. 10.5) are formed similarly to gold nanoparticles by a chemical reaction, which breaks down in an associating solvent (possibly also under the effect of additives) after a certain time. For the semiconductor nanoparticles, the band gap may depend on the particle size. This can be explained quantum mechanically with the particle in the box. So one can adjust the color via the particle size.
- There is a large variety of other inorganic nanoparticles, including rods (nanorods).
- The “Stöber particles” are formed by hydrolysis and subsequent recondensation of alkoxy silanes (Fig. 10.6). The process is triggered by basic pH. (In acidic environments, open networks form.) This class of reactions goes by the name "sol-gel synthesis." The process is robust and flexible (replacing silane with titanium, zirconium, etc. additives ...).
- Small particles in a gas phase can be produced cheaply by flame pyrolysis. For example, if $SiCl_4$ is blown through a flame in the presence of oxygen, SiO_2 particles are formed. However, these immediately aggregate (irreversibly) into larger clusters. A common product is called aerosil, also called "fumed silica".
  Small particles in a gas phase are called aerosols. Aerosol physics is much different from the physics of colloids.

---

[57] Kaolinite (discovered by Johann Friedrich Böttger while he was trying to turn urin into gold for his king, Augustus der Starke) is an essential component of porcelain. Today, it is mainly used for the production of high-gloss paper. Kaolinite induces blood clotting and is used in this capacity in wound dressing materials. Kaolinite also is used as a cheap filler in all types of polymers (and has more applications).

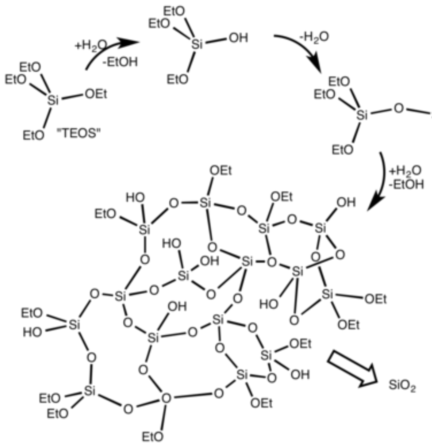


Fig. 10.6
On the synthesis of Stöber particles via the "sol-gel process".
en.wikipedia.org/wiki/Rummage_process

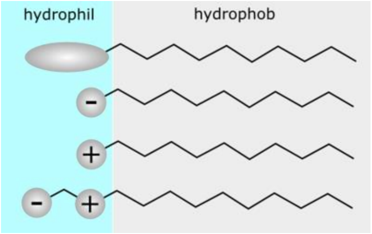


Fig. 10.7
Surfactants are often presented in "head-tail structure
en.wikipedia.org/wiki/Surfactant

10.2 Surfactants

Surfactants have already been discussed in the context of Gibbs adsorption isotherm, self-assembly, and association colloids. Here we collect a few properties from the view point of chemistry and applications.

– The English word is "surfactant" is derived from "surface active".

– Surfactants are not only amphiphilic, but also have strong lateral interaction between the hydrophobic alkyl chains. This is the hydrophobic interaction. Surfactants tend to self-assemble. A side remark on amphiphilicity: alkyl alcohols, for example, are amphiphilic. However, they do not form micelles. Many proteins are slightly amphiphilic. Many nanoparticles are slightly heterogeneous at the surface and therefore slightly amphiphilic, as well. Because the surface area per particle is large, nanoparticles often adsorb at the water/air and water/oil interfaces. Nanoparticles on the surfaces of oil droplets can stabilize these emulsions against aggregation in a purely geometric way. These are "Pickering emulsions".

– Functions of surfactants are
- lowering of the surface tension
- emulsifying
- act as wetting aids
- foaming (and also defoaming)
- stabilization of colloidal dispersions.

These functions can be combined with other functions. Polymerizable surfactants are called "*surfmers*" (composed of *surfactant* and *monomer*).

– For the representation one usually chooses the "head-tail-shape" (Fig. 10.7).
The head group can be
- anionic, e.g. sulfates, sulfonates, carboxylates, phosphates.
Sulfates are pH-stable. In the case of carboxylates, protonation occurs at low pH and the molecules

become neutral.
- cationic, e.g. quaternary amines ("quats"). These are pH stable, but expensive. Cationic surfactants are used a lot for hair care because hair (keratin) is often negatively charged. -zwitterionic, e.g. the betaines
- neutral ("non-ionic"), e.g. polyethylene oxide (PEO) or sugars. The PEO groups and the sugars are larger than the charged head groups, because otherwise sufficient amphiphilicity is not achieved.

- The hydrophobic part can consist of
  - alkyl chains, fluorinated chains, siloxanes and other groups
  - saturated or unsaturated chains. Unsaturated chains tend to crystallize less than saturated ones (a kink is formed). Therefore fluidity of the lamella is greater.
  - one or two chains: one chain for example in fatty acids, two chains for example in lipids. Two-chain surfactants are often more stable and form lamellae rather than vesicles because the shape parameter (Fig. 7.4) is closer to one.
  - Long or short chains. Long chains form more stable surfactants than short chains.
  - linear or branched chains. Branched chains do not form good lamellae (application: defoamers, wetting aids).

- Surfactants are sometimes classified according to the "hydrophilic-lipophilic balance". The smaller the "HLB value", the higher the lipophilic content. Initially, the HLB value was calculated from the number of hydrophilic and hydrophobic groups. Today, there are various other definitions. Table 10.1 shows applications for surfactants at the various HLB values.

- A molecular understanding of surfactant activity is mostly out of reach. Also for this reason (and because mixtures are often used), there is much empirical research.

- The production and in particular the purification of surfactants is costly. This is another reason why natural products and mixtures are often used.

10.3 Interfaces are often charged in water.

In water, almost all surfaces are charged. If the surface is not itself charged, there will be preferential adsorption of either H^+ or OH^- on the surface. Important charged surfaces and charged groups on surfaces are:

- Glass surfaces: Glass surfaces are negatively charged at neutral pH because the terminal SiOH groups are deprotonated.
- Other mineral surfaces: For Al_2O_3 and TiO_2, the charge depends on pH. These particles can be charged negatively or positively. The charge depends also on the crystallographic orientation of the surface. The isoelectric points of Al_2O_3 and TiO_2 are about 9 and about 6, respectively.
- Sulfates and sulfonates: Because the pK_A of sulfates and sulfonates is about 1, sulfates and sulfonates are pH stable. Sulfates and sulfonates are often parts of anionic surfactants (e.g. SDS, sodium dodecylsulfonate).

Table 10.1: HLB values and use

HLB value	**Usage**	**Miscibility with H O_2**
1.5 to 3	Defoamer	insoluble
3 to 8	for W/O emulsions	milky when stirring
7 to 9	Wetting agent	
8 to 18	for O/W emulsions	stable (milky) emulsion
13 to 15	Washing active substances	clear emulsion / clear solution
12 to 18	Solubilizer	clear emulsion / clear solution
en.wikipedia.org/wiki/HLB-value Hans-Dieter Dörfler, *Surface and Colloid Chemistry*, VCH, Weinheim, 1994, p. 198		

– Carboxyl groups: Because the pK_A is 3– 4, normal fatty acids are not pH stable. At low pH, they are protonated and thus neutral. That can be an advantage in case the charge state shall be controlled externally. Such surfaces often exist in “*responsive materials*”. Their properties “respond” to changes in pH. Carboxyl groups are contained in, for example, by polyacrylic acid (PAA).

– Amines: Amines are often positively charged. The positive charge of proteins (when present) is mainly due to amines. The quaternary amines ("quats", only these) are pH stable.

– Metal surfaces: Metal surfaces are often slightly negative in water because the quasi-free electron gas protrudes slightly further beyond the interface than the ions on the lattice.

– Adsorbed ionic surfactants: Surfactants readily adsorb to surfaces. When charged, they induce wettability by water.

– Neutral surfaces: Preferential adsorption of ions (e.g., more H^+ than OH^-) or a dipolar layer of an adsorbate also often occur on neutral surfaces.

The *isoelectric point* (IEP) is the pH, at which a surface (or also a protein) is neutral. To be distinguished from the IEP is the "point of zero charge". At this pH, colloidal particles carry no net charge. They may still be charged, locally. This situation occurs, for example, with layered silicates (positively charged at the edges, negatively charged on the surfaces).

Caution: Do not confuse the point of zero charge (a pH value) with the "potential of zero charge" in electrochemistry (an electrical potential).

Typical charges are on the order of a few $\mu C/cm^2$. A charge of 1 e/nm^{-2} corresponds to 16 $\mu C/cm^2$.

10.4 Side remark: cathodic electropainting

A targeted destabilization of colloids is used in the coating of car bodies. The car body is negatively charged. As OH^- is produced in the reaction $H_2O + e^- \rightarrow OH^- + 1/2\ H_2$, the electrolyte becomes alkaline. The dispersion is cationically stabilized with amines. Amines (with the exception of quaternary amines, that is, amines with all H-groups replaced by organic residues) are weak electrolytes. For instance, the equilibrium $R\text{-}NH_3^+ + OH^- \rightleftharpoons R\text{-}NH_2 + H_2O$ shifts to the righ-hand side. The groups become neutral, the particles are no longer collidally stable, and the dispersion precipitates.[58] The mechanism is shown in Fig. 10.8. An advantage of electrodeposition (compared to, for instance, spray coating) is that cavities are also coated. When the exposed parts of the body are covered with paint, the current flow shifts to the still uncoated areas because the paint is electrically insulating. The process is self-stabilizing.

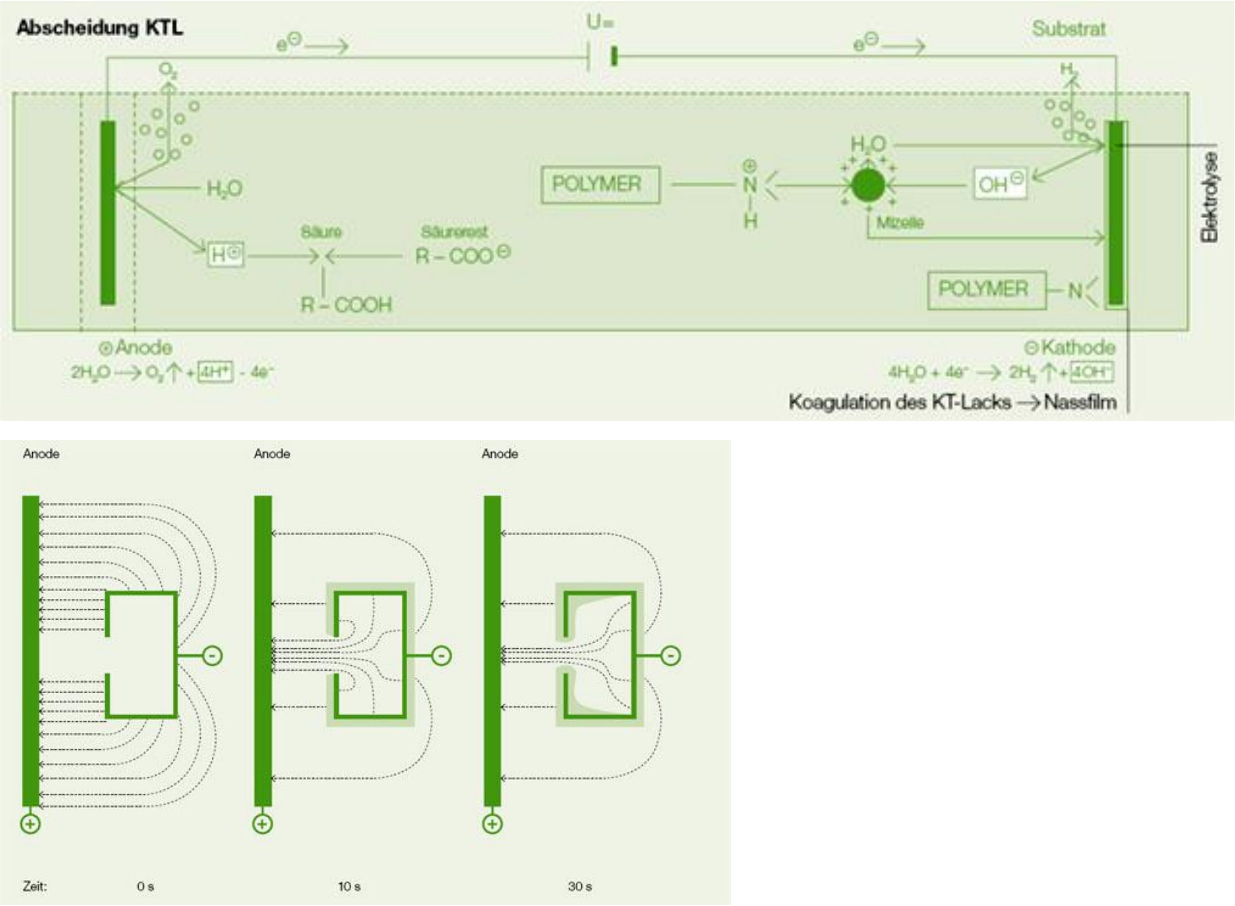


Fig. 10.8
Cathodic electrodeposition coating. The process The coating ensured of the surfaces inside cavities.
archiv.aktuelle-wochenschau.de/2006/week49b/week49b.html

---

[58] Stabilization with carboxylic acids (also a weak electrolyte, the charge of which depends on pH) would be cheaper, but the body of the car would then have to be operated as an anode. The pH would have to be pushed into the acidic range, meaning that H^+ would have to be formed in the reaction $H_2O \rightarrow 2H^+ + 2e^- + 1/2\ O_2$. Running a metal as an anode always runs the risk of corrosion by following reactions $M(s) \rightarrow M^+(aq) + e^-$ or $2M(s) \rightarrow M^{2+}(aq) + 2e^-$.

10.5 The Debye length, elements of the Debye-Hückel theory

We restrict ourselves to planar surfaces. On the way to the calculation of the Debye length, some elements of the Debye-Hückel theory will be explained.

The derivation starts from Maxwell's 1st law, which states that the divergence (the source density) of the electric field is proportional to the charge density.

Eq. 10.5.1

$$\nabla \cdot E = \frac{\rho}{\varepsilon\varepsilon_0}$$

Reminder: The divergence is defined as

Eq. 10.5.2

$$\nabla \cdot E = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}$$

We only consider planar surfaces. The field then is directed along *z*. All gradients occur along *z*. Maxwell’s 1st law turns into

Eq. 10.5.3

$$\frac{\partial E_z}{\partial z} = \frac{\rho}{\varepsilon\varepsilon_0}$$

We now use that the electric field is the negative gradient of an electric potential ψ :

Eq. 10.5.4

$$E_z = -\nabla\psi = -\frac{\partial\psi}{\partial z}$$

Inserting this relation into Eq. 10.5.3 results in

Eq. 10.5.5

$$\frac{\partial^2\psi}{\partial z^2} = -\frac{\rho}{\varepsilon\varepsilon_0}$$

This is the Poisson equation. For a given charge density distribution, the Poisson equation allows the electric field to be calculated. To do this, one must write a differential equation with source term (the right-hand side in Eq. 10.5.5) and - as always - with boundary conditions.

The local charge, in turn, is a function of the local electric potential. Using this dependence, Eq. 10.5.5 can be turned into a a closed expression only containing ψ. The charge is Boltzmann-distributed:

Eq. 10.5.6

$$\rho_\pm(z) = \pm e n_\pm \exp\left(\frac{\mp e\psi(z)}{k_B T}\right)$$

We consider 1-1 electrolytes in the following (ions with single charge). $n_\pm$ is the number density of ions in the limiting case of $z \to \infty$. Because of electroneutrality, n_+ and n_- (at large distance from the surface) are equal. We call this number density n_0 and write

Eq. 10.5.7

$$n_+ = n_- = n_0$$

Eq. 10.5.6 turns into

Eq. 10.5.8

$$\rho = \rho_- + \rho_+ = en_0 \left(-\exp\left(\frac{e\psi}{k_B T}\right) + \exp\left(\frac{-e\psi}{k_B T}\right) \right)$$

Inserting this result into Eq. 10.5.5 results in

Eq. 10.5.9

$$\frac{\partial^2 \psi}{\partial z^2} = -\frac{en_0}{\varepsilon\varepsilon_0} \left(-\exp\left(\frac{e\psi}{k_B T}\right) + \exp\left(\frac{-e\psi}{k_B T}\right) \right)$$

This is the *Poisson-Boltzmann equation*. It is a nonlinear partial differential equation in $\psi(z)$.[59]

The Poisson-Boltzmann equation is the starting point of Poisson-Boltzmann theory. The problem with the Poisson-Boltzmann equation is that it is nonlinear in ψ, because ψ occurs in the exponent. Therefore, one usually has to solve the equation numerically. This can be done. Poisson-Boltzmann theory is successful. Apart from the nonlinearity, Poisson-Boltzmann theory has similar weaknesses and strengths as Debye-Hückel theory, to which we will come in a moment.

To get from Poisson-Boltzmann theory to Debye-Hückel theory, one linearizes the Poisson-Boltzmann equation in ψ. One assumes $\varepsilon\psi/(k_B T) << 1$ and Taylor-expands the exponential to first order:

Eq. 10.5.10

$$\exp\left(\frac{e\psi}{k_B T}\right) \approx 1 + \frac{e\psi}{k_B T}$$

Inserting this result into Eq. 10.5.9 leads to

Eq. 10.5.11

$$\frac{\partial^2 \psi}{\partial z^2} = \frac{en_0}{\varepsilon\varepsilon_0} \left[\left(1 + \frac{e\psi}{k_B T}\right) - \left(1 - \frac{e\psi}{k_B T}\right) \right] = \frac{2n_0}{\varepsilon\varepsilon_0} \frac{e^2\psi}{k_B T}$$

This is the desired linear differential equation in ψ.

---

[59] Strictly speaking, Eq. 10.5.6 is an ordinary differential equation (in 1D). It turns into a partical differential equation in 3D. ψ then is a function x, y, and z. One writes $\nabla^2 = (\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2)$ on the left-hand side.

Having obtained this equation from a Taylor expansion, one might suspect that it only is valid for potentials much below k_BT/e. The value of k_BT/e at room temperature is 25.7 mV.[60] However, favorable coincidences cause the Debye-Hückel theory to give good results up to voltages of about 50 mV.

We attempt to solve Eq. 10.5.11 with the ansatz

Eq. 10.5.12

$$\psi(z) = \psi_\delta \exp\left(-\frac{z}{r_D}\right)$$

ψ_δ is the potential at the bottom of the diffuse double layer (at the outer edge of the outer Helmholtz layer). r_D is the Debye length. Inserting into Eq. 10.5.11 leads to

Eq. 10.5.13

$$r_D = \sqrt{\frac{\varepsilon\varepsilon_0 k_B T}{2n_0 e^2}}$$

Without proof, we provide the extension of this formula to multivalent ions, which is

Eq. 10.5.14

$$r_D = \sqrt{\frac{\varepsilon\varepsilon_0 RT}{2F^2 I}}$$

The ionic strength I is defined as

Eq. 10.5.15

$$I = \frac{1}{2}\sum_i z_i^2 c_i$$

c_i is the concentration in units of mol/m^3. Note: The unit is mol/m^3, as opposed to mol/L. Otherwise a factor of 1000 would be needed to convert between liters and m^3. For 1-1 electrolytes, the ionic strength is the same as the salt concentration. As an example, consider NaCl:

Eq. 10.5.16

$$I = \frac{1}{2}\left(c_{Na^+} + c_{Cl^-}\right) = c_{NaCl}$$

For $CaCl_2$, the ionic strength is larger than the salt concentration:

Eq. 10.5.17

$$I = \frac{1}{2}\left(4c_{Ca} + c_{Cl^-}\right) = \frac{1}{2}\left(4c_{CaCl_2} + 2c_{CaCl_2}\right) = 3c_{CaCl_2}$$

---

[60] The term k_BT/e also appears in the Nernst equation. There it is usually written as RT/F with $R = N_A k_B$ the gas constants and $F = N_A e$ the Faraday constants.

For water at room temperature, one may remember the relation

Eq. 10.5.18

$$r_D \approx \frac{0.3\,\mathrm{nm}}{\sqrt{I\,[\mathrm{M}]}}$$

The ionic strength here has units of mol/L (as opposed to mol/m^3, cf. Eq. 10.5.15). Remember: Adding salt lowers the Debye length (Fig. 10.11).

In pure water at pH 7, H$^+$ and OH$^-$ are the only ions contributing electrostatic screening. The Debye length in pure water is about 1 µm. In biofluids, the ionic strength is about 100 mM. The Debye length then is about 1 nm. When the Debye length is about as large as molecules themselves, the Debye-Hückel theory must reach its limits, and it does. Our limited understanding of electrostatic interactions in water is a fundamental problem in molecular biophysics.

Without proof we extend Eq. 10.5.12 to three dimensions:

Eq. 10.5.19

$$\psi = \frac{-e^2}{4\pi\varepsilon\varepsilon_0}\frac{1}{r}\exp\left(-\frac{r}{r_D}\right)$$

To check for correctness, insert Eq. 10.5.11 and remember that the radial part of the ∇^2-operator in spherical coordinates is given as $1/r^2\partial_r\,(r^2\partial_r)$

We calculate the Debye length a second time, avoiding differential equations. This second derivation also clarifies that the thickness of the diffuse double layer results from an interplay between entropy and enthalpy. We view of the double layer as a *capacitor*. This capacitor wants to minimize its electrostatic energy. At the same time, the counterions forming one plate of the capacitor create an osmotic pressure. Balancing the force of electrostatics against the osmotic pressure takes care of the balance between energy and entropy.

The electrostatic energy contained in a capacitor is given as $E = CU^2/2$ with $C = \varepsilon\varepsilon_0 A/d_C$ the capacitance, A the area, ε the dielectric constant, $\varepsilon_0 = 8.8\ 10^{-12}$ C/(Vm) the permittivity of the vacuum, d_C the distance between the two plates, and U the voltage. Note that the voltage is a fixed boundary condition here (as opposed to the charge). (The same problem exists in the other versions of the Debye-Hückel theory).

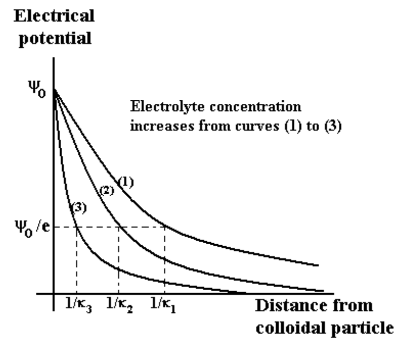


Fig. 10.9
Addition of salt lowers the Debye length, thereby deteriorating colloid stability.
www.colloid.ch/index.php?name=dlvo

A side remark: The double layer capacitance is often measured and interpreted. Typical values are some 10's of µF/cm^2.

There electrostatic attraction causes a pressure, given as

Eq. 10.5.20

$$p_{es} = -\frac{\mathrm{d}U_{es}}{\mathrm{d}V} = -\frac{1}{A}\frac{\mathrm{d}}{\mathrm{d}d_C}\left(\frac{1}{2}CU^2\right) = -\frac{1}{A}\frac{\mathrm{d}}{\mathrm{d}d_C}\left(\frac{1}{2}\frac{A\varepsilon\varepsilon_0}{d_C}U^2\right) = \frac{1}{2}\frac{\varepsilon\varepsilon_0 U^2}{d_C^2}$$

The subscript *es* stands for “electrostatic”.

The net entropic pressure is the *difference between* the osmotic pressures inside the capacitor and outside the capacitor (that is, in the bulk). The osmotic pressure inside the capacitor is

Eq. 10.5.21

$$p_{osm} = RT\frac{n_B}{V} = RT\left(c_+ + c_-\right) = N_A k_B T\left(c_+ + c_-\right) = k_B T\left(n_+ + n_-\right)$$

$$n_\pm = n_0 \exp\left(-\frac{eU}{k_B T}\right) \approx n_0\left(1 \mp \frac{eU}{k_B T} + \frac{1}{2}\left(\frac{eU}{k_B T}\right)^2\right)$$

Note that $n_\pm$ must be Taylor-expanded to second order. The osmotic pressure acting onto the plate from the inside thus is

Eq. 10.5.22

$$p_{osm} \approx k_B T n_0\left(2 + \left(\frac{eU}{k_B T}\right)^2\right)$$

Subtracting the pressure acting from the outside ($2RTc_0 = 2k_BTn_0$) leads to

Eq. 10.5.23

$$\Delta p_{osm} \approx k_B T n_0\left(\frac{eU}{k_B T}\right)^2$$

We assume $p_{es} = \Delta p_{osm}$ and obtain

Eq. 10.5.24

$$\frac{U^2}{2}\frac{\varepsilon\varepsilon_0}{d_C^2} \approx k_B T n_0\left(\frac{eU}{k_B T}\right)^2$$

Solving for d_C leads to

Eq. 10.5.25

$$d_C^2 \approx \frac{\varepsilon\varepsilon_0 k_B T}{2e^2 n_0}$$

This thickness should be compared to the Debye length. It actually agrees with the Debye length, which is fortuitous.

Remember

- The diffuse double layer is the result of and interplay between entropy and enthalpy.
- The thickness of the diffuse double layer is the Debye length. In water at 298 K, $r_D \approx 0.3$ nm/(I[mol/L]$^{1/2}$).

10.6 Dielectric and electrostatic screening

Central to the behavior of ions in water are dielectric screening and electrostatic screening. Dielectric screening is caused by the dipole moment of the water molecules. When an electric field with a direction normal to an air-water interface intersects that interface, the field in water is weaker than in air by a factor of 78 (where 78 is the dielectric constant, ε). The orientated dipoles at the surface screen the field a large amount. The hydration shells around the ions screen the field of the ions by the same factor of 78. Dielectric screening allows to dissolve ions in water. Dielectric screening leads to the hydration enthalpy. Because the hydration enthalpy is comparable in magnitude to the lattice enthalpy of the salt, salt dissolve in water.

The electrostatic screening is caused by the counterion cloud. The counterion cloud is an increased probability density to find a counterion in the vicinity of some ion. Remember: An ion cannot be surrounded by cloud of many counterions because that would violate electroneutrality. The probability density hints as a weakness of Debye-Hückel. Replacing the counter-ions by their probability density does not always catch the entire physics This substitution is typical for “mean-field theories”. Not covered by these mean-field theories are the "correlations". Correlations in this context are effects which are not captured by the mean-field theory. The neglect of ion correlations is one of a few shortcomings of the Debye-Hückel theory.

The difference between dielectric and electrostatic screening can be remembered by the equation for the electrical potential around an ions, which is:

Eq. 10.6.1

$$\psi = \frac{-e^2}{4\pi\varepsilon_0 r}\frac{1}{\varepsilon}\exp\left(-\frac{r}{r_D}\right)$$

The first term is the Coulomb potential, which also acts in a vacuum. Dielectric screening weakens this potential by the factor 1/ε. The exponential term quantifies the electrostatic screening. r_D is about the radius of the counterion cloud. Because of electrostatic screening, there can never be large-scale electric fields in equilibrium situations in water. These field exist nevertheless, if

- they are alternating fields whose frequency is so high that the formation of the counterion cloud cannot follow the field, or
- the field in the electrolyte causes an electric current into the electrodes. The field then follows from Ohm’s law ($U = RI$ with U the voltage, I the current and R the resistance).

The first situation is exploited in measurements of the conductivity of electrolytes. An alternating voltage is used. The frequency is in the kHz range. The second situation occurs in electrophoresis.

Locally (on the scale of Debye length and below), electric fields do exist in water. They can even be much larger than the fields known from the macroscopic world. This is so because the very strong fields cause electric breakdown (e.g., an arc) on macroscopic scales. Electric breakdown does not occur on the scale of the Debye length. On the scale of Debye length, electroneutrality may be violated.

Two side remarks:

- There is an analog to the Debye length in plasma physics and in the physics of semiconductors.[61]
- In metals, the “counterions” of the electrons (the ions on the lattice) are immobile. Because of that, electrostatic screening does not occur and electric waves travel through the wire i at almost the speed of light.

---

[61] https://de.wikipedia.org/wiki/Debye-Länge

The existence of the counterion cloud is a consequence of the *interplay of enthalpy and entropy*. If there was no drive to large entropy, one counterion would condense onto each ion. Indeed, there is such a form of counterion *condensation*. Counterions are often tightly bound to the central ion or a charged surface. However, because of entropy, some counterions remain in a mobile state near the central ion or the surface.

Before the value of the Debye length is derived, a remark on the strength of the electrostatic interaction: This is enormous. It is not perceived as enormous in everyday life only *because* it is screened. As an example, calculate the force between two spheres at a distance of 1 m, each containing 1 mol protons and 1 mol electrons (uncompensated). Consider also that the "strong interaction" which acts between the nucleons in atomic nuclei can hold a maximum of 82 protons in the atomic nucleus in the long run (whereas uranium with the nuclear charge number 92 is admittedly also almost stable).

<u>Remember</u>

- In water, the electrostatic interaction is screened in two ways. The polarizability of water lowers all fields by a factor of about 78 (by the static dielectric constant). The counter ion cloud additionally leads to an exponential decrease of the potential. The potential is given as
  $$\psi = \frac{-e^2}{4\pi\varepsilon_0 r}\frac{1}{\varepsilon}\exp\left(-\frac{r}{r_D}\right)$$

10.7 Organization of counterions at charged surfaces

The conditions at charged surfaces are rather complicated. Although the separation between the different "layers" (Fig. 10.9) is not always clear-cut, it is possible to distinguish between different forms of attachment of counterions to charged interfaces. We progress from the outside to the inside:

- Due to entropy, counter ions usually do not attach to the surface to the extent that the charge of the surface would be fully compensated by the adsorbed ions. A minority of ions remain at a certain distance from the surface and form the *diffuse double layer*. The thickness of this diffuse double layer (the Debye length) results from a balance between entropy and enthalpy. The thickness of the diffuse double layer depends on the salt concentration. Typical values are in the range of a few nanometers. The Debye length is discussed in Sect. 10.5.[62]

- Some ions attach to the surface *without* stripping the hydration shell. These form the *outer Helmholtz layer*.

- Some ions attach to the surface, shedding the hydration shell. These form the inner Helmholtz layer. One also speaks of "adsorbed ions". The evidence for the difference between the outer and inner Helmholtz layer comes firstly from the kinetics of electrochemical reactions and secondly from studying the dependence of the double layer capacitance on the potential of the surface. The stripping of the hydration shell accounts for a significant contribution to the activation energy of electrochemical reactions. We pass over the effects of the two Helmholtz layers on the double layer capacitance.

Details of the potential profile at charged surfaces are fundamentally a difficult question. The measurement of the "surface potential" is fraught with the same conceptual and technical difficulties.

---

[62] Do not confuse the diffuse double layer with the Nernst layer. The latter is many 100 µm thick. The Nernst layer is the unstirred layer within which all transport is diffusive.

"Surface potentials" are usually determined "electrokinetically". For particles this means the determination of the drift velocity v_{drift} in the electric field E ("electrophoresis"). For planar surfaces one determines the streaming potential. In both cases there is a relation between the velocity, on the one hand, and the ζ potential (zeta potential), on the other. We only provide the equation adapted to electrophoresis (in its most simple form[63]), which is

Eq. 10.7.1

$$\frac{v_{drift}}{E} = \mu_{ep} \approx \frac{\zeta\varepsilon\varepsilon_0}{\eta}$$

μ_{ep} the electrophoretic mobility. Note: For colloids, the mobility depends weakly on the radius. (This is different for proteins.) ζ is the ζ potential. This raises the question of where on the potential-distance curve the ζ potential is located (Fig. 10.10). The ζ potential is located at the "plane-of-zero-shear". At a certain distance from the surface, the distance-velocity relation is linear. (The "velocity" v here is the tangential velocity of the fluid relative to the surface). Extrapolating this relation to v = 0 yields the plane-of-zero-shear (also: "slip plane", see Box 10.1).

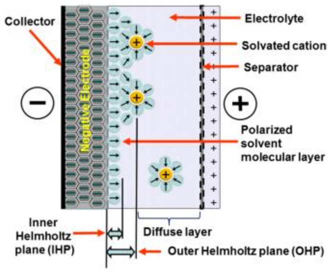


Fig. 10.10
Organization of counterions on charged surfaces. A distinction is made between the inner Helmholtz layer, the outer Helmholtz layer and the diffuse double layer.
en.wikipedia.org/wiki/Double-layer_capacitance

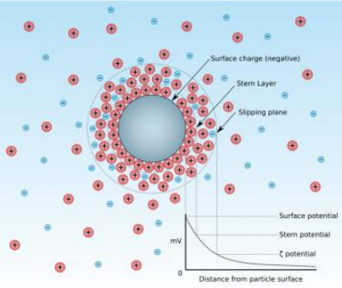


Fig. 10.11
The ζ potential is the electrical potential at the slip plane
en.wikipedia.org/wiki/Zeta_potential

Two remarks:

- For planar surfaces, the ζ potential is inferred from the "streaming potential" or the "electroosmotic flow". Both are "electrokinetic phenomena", similar to electrophoresis.
- The ζ potential can also be determined by electroacoustic means. If there is a density difference between a particle and medium, a sound wave causes a relative motion and consequently a radio wave.

<u>Remember</u>

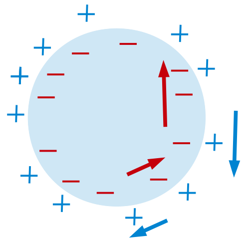
- On charged surfaces, a distinction is made between the inner Helmholtz layer, the outer Helmholtz layer and the diffuse double layer.

---

[63] For very small particles, the relation $\mu \approx q/(6\pi\eta r_H)$ is more suitable. q is the charge and r_H is the hydrodynamic radius.

- The ζ potential is the potential at the plane-of-zero-shear.

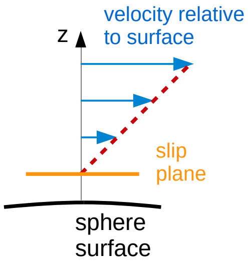
Box 10.1 Electrophoresis in a nutshell



The following remarks concern the electrophoresis of particles (as opposed to polymer chains). The motion of charged particles in an electric field is part of "electrokinetics" (as opposed to electrochemistry). Other electrokinetic phenomena are the electroosmotic flow, the streaming potential and sedimentation potential. All these phenomena depend on the ζ-potential (zeta-potential).

From a distance, charged particles in an electrolyte appear as neutral due to the counter-ion cloud (synonymous to the double layer). They still move in an electric field because the double layer is sheared off the particle by the electric field. (top figure on the right-hand side). The particle and its counter ion cloud move into opposite directions. The simplest model of electrophoresis predicts the drift velocity, v_{drift}, to be

$$v_{drift} = \mu_{ep} E \approx \frac{\zeta\varepsilon\varepsilon_0}{\eta} E$$



μ_{ep} is the electrophoretic mobility, E is the electric field, ε is the dielectric constant ζ is the ζ-potential and η is the viscosity. Note that the drift velocity in this very simple does not depend on the size of the particle. Neither does it depend on the charge. It depends on the electric potential at the slip plan (lower figure). Because the ζ-potential governs the electrophoretic mobility, it is of much practical importance. It is easily measured.

The slip plane is the distance from the surface, where velocity profile extrapolated to v = 0 hits zero. It usually is located slightly above the surface because the viscosity is slightly increased close to the surface. (It can also lie below the surface, in case the viscosity is decreased at the surface. In this case the – negative – distance to the surface is called "slip" length.) The location of the slip plane is not easily predicted from simple models. Because hydrodynamics enters the definition of the ζ-potential, differences is ζ between particles can have a few different reasons.

10.8 Ion-specific effects and the Hofmeister series

The counterion cloud is the central prediction of Debye-Hückel theory. Should the Debye-Hückel theory be in applicable, the counterion still are the starting point for all further discussions. Below is an (incomplete) list of phenomena covered by Debye-Hückel theory:

- The forces between charged surfaces
  We had calculated above only the course of the electric potential. The force-distance law, which is derived from the electrostatic part of the DLVO theory (section 10.9) also decays exponentially with distance. The characteristic length is the Debye length.

- The critical salt concentration at which electrostatic stabilization breaks down
  Why salt has an unfavorable effect on colloid stability is discussed in Sect. 10.9. This critical concentration goes by the name "critical coagulation concentration", ccc). According to Debye-Hückel theory, it depends on the charge number of the ions, but not on their chemical nature. The higher the charge, the larger the tendency to salting out (the latter content of the Schulze-Hardy rule).

- The activity coefficients of ions in water
  At small concentrations, Debye's limiting law applies: $\log_{10}\gamma_{\pm} = -A|z_+ z_-|\sqrt{I}$ with $A = 0.509\ (\text{L/mol})^{1/2}$ for water at $T = 298$ K, $z_{\pm}$ the charge numbers, and I the ionic strength.

- The mobility of ions in electric field (Kohlrausch's square root law)
  The counter ion cloud delays the transport of the central ion. According to Debye-Onsager theory, there is an "electrophoretic effect" and a "relaxation effect". Both scale with the square root of the ionic strength (like the Debye length).
- The rate constants of reactions in solution and in particular their dependence on the salt concentration ("kinetic foreign salt effect").

Shortcomings of the Debye-Hückel theory include:

- Debye-Hückel theory is based on a Taylor expansion in the voltage and should therefore only be valid for voltages below 25.7 mV (25.7 mV = $k_BT/e = RT/F$ at T = 298.15 K). By coincidence, different errors cancel. Debye-Hückel theory often is applied at potentials larger than 25 mV.
- Ion correlations: Debye-Hückel theory is a "mean-field theory" that replaces the individual ions with an averaged ion density.
- Finite ion radii
  Debye-Hückel theory must fail when the Debye length approaches the ion radius (including the hydration shell). Beyond concentrations of 100 mM, deviations can be expected. In the case of Debye's limit law, this can be seen, for example, in Fig. 10.12
- Van-der-Waals attraction
- H-bonds, hydration shell

Debye-Hückel theory is universal in the sense that the chemical properties of an ion – apart from the charge – play no role. If this were correct, a "sodium-potassium asymmetry" (the energy source of active signal transduction in nervous tissue) would remain without consequences. Those effects, that depend on the chemical nature of the ions go by the name "ion-specific effects" (ISEs). Among the ion-specific effects are

- Deviation of the activity coefficient from Debye's limit law (Fig. 10.12)
- The influence of ions on solubility, reactivities (also of enzymes), and the CMC
- The influence of ions on the surface energy.

ISEs are often organized along the Hofmeister series. In 1888, Hofmeister determined the critical salt concentration at which a certain protein (especially albumin from the hen's egg) precipitated and ranked the salts according to this concentration. Today, this concentration would be called the "critical coagulation concentration (ccc)." Debye-Hückel theory makes a prediction for the ccc. As Hofmeister showed, it does not agree well with experiment. In particular, he also observed "salting-in", not only a salting-out as predicted by the Debye-Hückel theory. A similar ordering as observed by Hofmeister is also found in many other ISEs. Caution is advised. Sometimes the direction by which the ions are ordered flips when concentrations are changed.

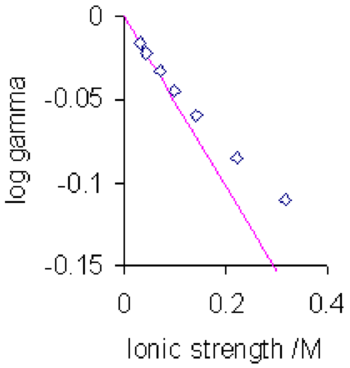


Fig. 10.12
Beyond an ionic strength of about 100 mM, many deviations from the Debye-Hückel theory occur. Shown here is log $\gamma_\pm$ of KBr in water at 25°C (with $\gamma\pm$ the activity coefficients). The line is the prediction of Debye's limit law.
en.wikipedia.org/wiki/Debye-Hückel_theory

Hofmeister attributed the ISEs to the influence of the ions on the structure of the surrounding water. Ions which disturb this structure he called chaotropic; ions which stabilize it he called cosmotropic. Although this view is not strictly equal to the current understanding, some of the effects can be remembered based this idea.

Remember

- The Debye-Hückel theory is the starting point for the discussion of electrostatic effects in water. It is universal.
- Debye-Hückel theory does not cover the ion-specific effects. These can arranged along the Hofmeister series.

10.9 Colloid stabilization according to DLVO theory, steric stabilization

In the context of DLVO theory, the interactions between colloidal particles are assumed to be the sum of van-der-Waals attraction and electrostatic repulsion. The electrostatic repulsion is described as an exponential, following Debye-Hückel theory. (We omit the derivation of the interaction potential from the electric potential). The interaction potential is

Eq. 10.9.1

$$V_{rep} \approx V_{\delta} \exp\left(\frac{-D}{r_D}\right)$$

V_δ is a prefactor. δ stands for the upper edge of the outer Helmholtz-plane, which at the same time is the lower edge of the diffuse double layer. D is the distance between the surfaces of the two spheres (as opposed to the distance between the sphere centers). DLVO theory in its simple form applies to distances much smaller than the sphere radius a (applies when $D \ll a$). This condition is quite realistic for technical polymer dispersions, where high solids contents are used. Particles are packed as densely as possible while still allowing them to flow. (A side remark: Flowability at high solids content is enhanced by wide particle size distributions, because the small particles fit into the interstices between the large particles).

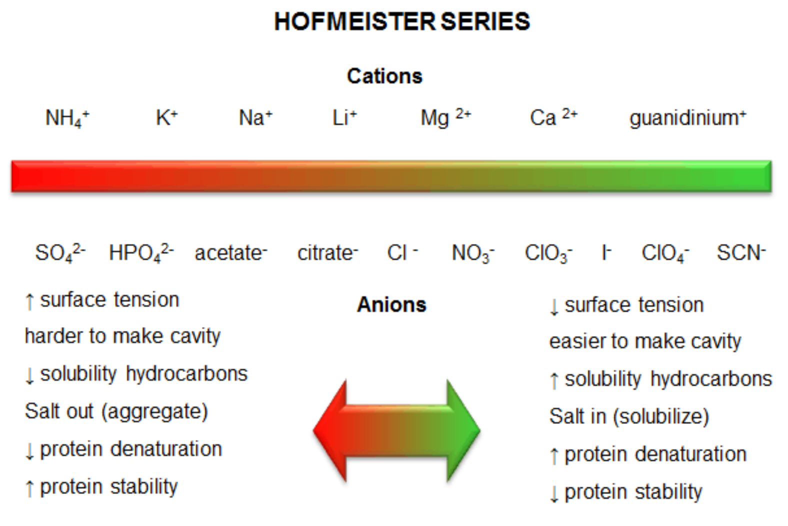


Fig. 10.13
Properties of ions according to their position in the Hofmeister series
employees.csbsju.edu/hjakubowski/classes/ch331/protstructure/hofmeister.gif

The van-der-Waals attraction is given as

Eq. 10.9.2

$$V_{att} \approx -\frac{Aa}{12D}$$

A is the Hamaker constant.[64] Typical values are around of 10^{-20} J. Note that the vdW attraction increases with particle size, while repulsion does not (Eq. 10.9.1). Large particles aggregate more easily than small particles.

A side remark: Eq. 10.9.2 contains the term $1/D$ (as opposed to $1/D^6$), although the distance dependence of the vdW interaction between two molecules scales as $1/r^6$. For colloidal spheres, an inverse sixth power results when the sphere distance is much larger than the sphere radius (when $D >> a$). Here, we consider the opposite case ($D << a$). Many different molecules contribute to the attraction. One has to integrate and in the course of these integrations a r^{-6}-law becomes a D^{-1}-law

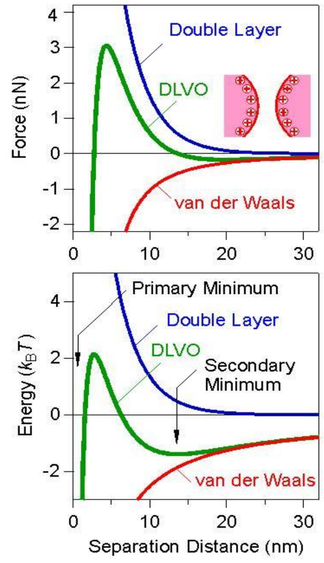


Fig. 10.14
The interaction between two colloidal spheres according to DLVO theory.
www.colloid.ch/index.php?name=dlvo

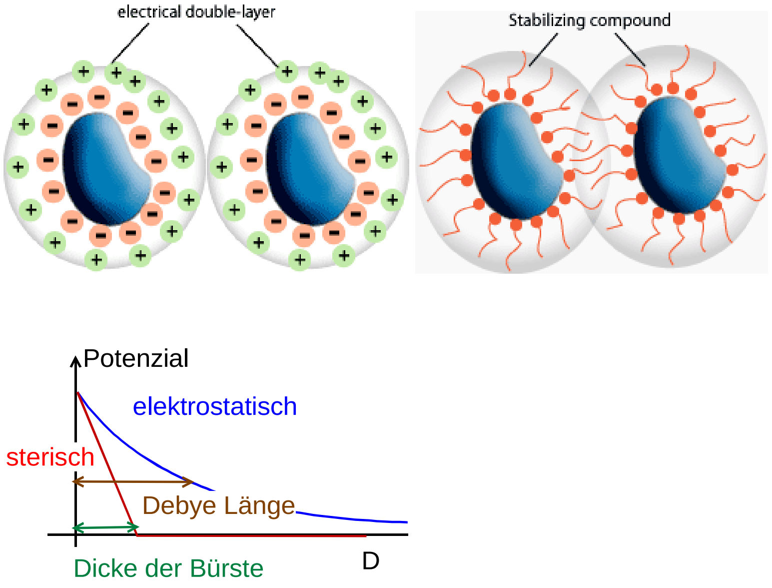


Fig. 10.15
On electrostatic and steric stabilization
www.inkline.gr/inkjet/newtech/tech/dispersion/

As Fig. 10.14 illustrates, the sum of electrostatic repulsion and vdW attraction results in an aggregation barrier if - first - the prefactor before the repulsive term is large enough and if - second - the range of the repulsion is large enough. This also names the two central failure mechanisms of electrostatic stabilization. For example, the prefactor becomes small when charge on the particles is lost due to lack of pH stability. The range also becomes small when, after the addition of salt, the Debye length becomes small.

For the two reasons mentioned above, electrostatic colloid stabilization is not very robust. One can help oneself by additionally (or alternatively) applying *steric* stabilization (Fig. 10.15). Steric

---

[64] This is an effective Hamaker constant. Write $A = (A_{11}^{1/2} – A_{22}^{1/2})^2$ with A_{11} and A_{22} the Hamaker constants of the particles and the liquid. Note that the interaction is attractive even for the case $A_{11} < A_{22}$. Two air bubbles in a liquid attract. There can be a net vdW repulsion if the two "particles" have different Hamaker constants. For this case, there is an even more complicated compound Hamaker constant that accounts for three different materials. This latter effective Hamaker constant can be negative. An air bubble and a solid particle in water can repel each other (due to the vdW interaction in connection with Archimedes' principle).
Basically, the Hamaker constant has the unit of one energy. It can be divided by k_BT. The Hamaker constant is usually slightly larger than k_BT. Note, however, that the Hamaker constant rarely competes directly with the thermal energy. There are almost always geometry factors still involved, including (Eq. 10.9.2).

stabilization is achieved by polymer brushes. Polymer brushes here mean terminal anchored linear chains in a good solvent. Two polymer brushes always repel each other for entropic reasons. The preparation of brushes on latex particles is relatively simple for acrylate polymers: acrylic acid is added to the formulation at a concentration of about 1%. These relevant partial chains are water-soluble and therefore form the required brush-like structures on the surface. Non-ionic surfactants with their relatively long hydrophobic moieties also contribute to steric stabilization.

It should be noted in the margin a subtlety in the DLVO potential. The sum of repulsive and attractive potential leads to a flat secondary minimum at large distances. This minimum is usually of minor importance, but can have consequences. It can lead to flocculation (Sect. 10.10).

As far as quantitative details are concerned, the DLVO theory must have at least the same weaknesses as the Debye-Hückel theory. However, it explains some essential elements of colloid stability (and how it can be lost) well.

Remember

- The electrostatic repulsion creates an aggregation barrier (if the charge and the Debye length are large enough).
- Electrostatic colloid stabilization suffers from the addition of salt.
- "Steric stabilization" with polymer brushes is a second mechanism for colloid stabilization.

10.10 Flocculation

Flocculation refers to a reversible accumulation of particles on each other. In contrast, "aggregation" is irreversible. Aggregation means aggregation into the primary minimum (Fig. 10.14). Dominant here is the vdW interaction. Aggregation is difficult to break up even by ultrasound and/or strong shear fluxes. Flocculation is already dissolved by moderate shear forces (e.g. generated by dipping a brush into the paint).

Flocculation is often desired because

- it has a favorable influence on the rheological properties during processing. Here, the shear thinning is to be mentioned above all
- a flowing of the film on the wall ("sagging") is avoided
- sedimentation of heavy pigments is hindered.

There are several mechanisms that lead to flocculation:

- A preferential interparticle spacing in the secondary minimum of the DLVO curve (Fig. 10.14).
- Bridging flocculation (Fig. 10.16) according to This mechanism requires weakly adsorbing chains.
- Depletion flocculation (Fig. 10.17). Depletion flocculation is caused by a bimodal (or multimodal) size distribution. If the smaller species does not fit into the gap between the larger particles, an osmotic vacuum is created in the gap.

Flocculation produces weak gels. In the terminology of rheology, these substances are called "Bingham fluids". They flow after a certain critical stress (the yield-stress) is exceeded. Such weak gels are also very common among foods (ketchup, mustard, ...). Pastes have similar properties.

Fig. 10.16
Mechanism of bridging flocculation
community.dur.ac. uk/ sharon.cooper/lectures /colloids/interfacesweb1 .html

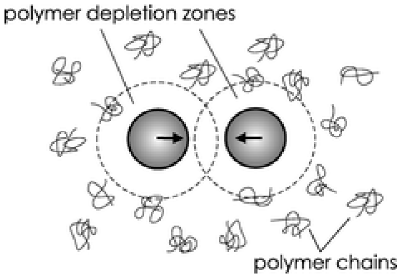


Fig. 10.17
Mechanism of depletion flocculation
Fan, T.-H.; Tuinier, R., Hydrodynamic interaction of two colloids in non-adsorbing polymer solutions. Soft Matter ***2010,*** *6, 647.*

11 Film formation from latex dispersions

11.1 General

- *Literature on film formation:*
  - Keddie, J. L.; Routh, A. F., *Fundamentals of Latex Film Formation: Processes and Properties*. Springer: 2010.
- Mischke, P.; Brown, R., *Film Formation in Modern Paint Systems*. Vincentz Network: 2010.

Film formation in a way reverses the process of colloid generation. It is a controlled solidification by drying. In the following, we are concerned with coatings (as opposed to sintered bodies). We are not concerned with molecularly thin coatings. Typical thicknesses are 100 µm. (In automotive engineering: more likely 150 µm). This thickness allows to cover dust grains and to compensate a certain roughness. The material can dissipate mechanical energy to achieve some impact resistance. Some abrasion can be tolerated over the life of the paint.

A very important aspect of film formation is that polymers can relieve (relax) shrinkage-induced stresses. This is a special property of polymers whose glass transition temperature is not far above the working temperature. This is the role of polymers as *binders*.[65] The binder may well be present in such small quantities between the "pigments" that the "film" is porous and the "binder" only forms bridges between the particles. This is the case, for example, with interior wall paints. These would have to be permeable to moisture. With the exterior wall, this is different.

Example applications include:
- Protective coatings (automotive industry, shipbuilding, architecture, wood preservation...)
- Adhesives[66]
- Rubber gloves
- Back coatings for carpets (a large but declining market).
- the most diverse types of inks (these usually not applied planar).[67] The "pigments" can be functional carriers of all kinds (photovoltaic, sensory, ...).
- Electrodes in lithium batteries consist of the "active", graphite for electrical conductivity and the polymer PVDF. They are applied as a film and then rolled.
- Auxiliary materials for paper production

The production of latex-based paints has a number of advantages. Because mainly water evaporates during drying (instead of organic solvents), latex-based paints are considered environmentally friendly.[68] Unfortunately, latex dispersions still contain "film-forming aids." These are organic molecules. You can smell these substances when water-based paints dry. Furthermore, latex dispersions can have a high solids content while maintaining moderate viscosity. Solution-based dispersions with a concentration of 50% or more (as is common with latex dispersions) are no longer

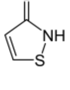
---

[65] The "CSH phases" in concrete play a comparable role. CSH stands for calcium silicate hydrate. These are fine and, within certain limits, flexible lamellae whose exact structure is not well known.

[66] With adhesives, stress relaxation is not a problem because the materials are soft.

[67] 3D printing is not considered in the following. In polymer 3D printing, the most common method is fused-deposition molding (FDM). There are other methods. FDM is based on cooling liquid polymer. No drying takes place.

[68] Consider, however, that water-based processes (always) require the use of biocides. Biocides are playing an increasingly important role as water-based processes become more important. There is a complicated set of regulations. Below is the structure of isothiazolinone, an important component.



flowable, but latex dispersions are. Third, emulsion polymerization is a relatively inexpensive manufacturing process available. Caution: Do not confuse emulsion polymerization with miniemulsion polymerization.

11.2 The film formation process

The "film formation process" is commonly divided into three phases (Fig. 11.2):

- Evaporation of the water ("stage I")
- Particle deformation ("stage II")
- Polymer interdiffusion ("stage II").

All three phases have their problems and open specific possibilities to adjust the film properties.

Box 11.1

Emulsion polymerization

Emulsion polymerization does not produce dispersed polymer particles from emulsified monomer droplets in the course of polymerization. This latter process is at the core of miniemulsion polymerization. An essential aspect of miniemulsion polymerization is the stabilization of the emulsion against Ostwald ripening with a "hydrophobe" (section 3.5).

In emulsion polymerization, one starts with relatively large monomer droplets, which, however, only play the role of a monomer reservoir. The monomers must be readily soluble in water. The initiator is also water-soluble. A large number of micelles are present. Initially these are empty, later they swell with monomer. (Much later, these micelles become the polymer particles.) The initiator breaks down into two radicals in the water phase. After the monomer is easily soluble in water, oligomers are formed. Due to their size, oligomers are somewhat less soluble in water than the monomers. Therefore, sooner or later they enter the interior of the micelles. There, polymerization then proceeds further. In principle, the growing chains can of course also enter the (large) monomer droplets and form a chain there. However, because there are far fewer monomer droplets than micelles, the latter process is rare.

A certain disadvantage of emulsion polymerization is that it cannot be carried out continuously. It is a batch process. Attempts have been made to build appropriate tubular reactors. Reactor fouling – the risk of a solid phase forming – remained a problem.

A complication in emulsion polymerization can be that new particles are formed in a late phase. In seeded emulsion polymerization (also known as semi-batch polymerization), small particles are introduced and the particle nucleation phase is separated from the growth phase.

In the "semi-batch process", not all reagents are present from the beginning. Rather, monomer, initiator and possibly other reagents are added from above during the reaction. In this way, one can counteract "compositional drift" in copolymerization, for example. One can also produce "core-shell particles" (e.g. with a hard core and a soft shell) by changing the monomer during polymerization. This creates "nanocomposite" particles. Nanocomposites, unlike copolymers, exhibit some spatial heterogeneity. Nanocomposites are also important because "latex blends" tend to segregate due to the small number of particles and therefore also the small mixing entropy.

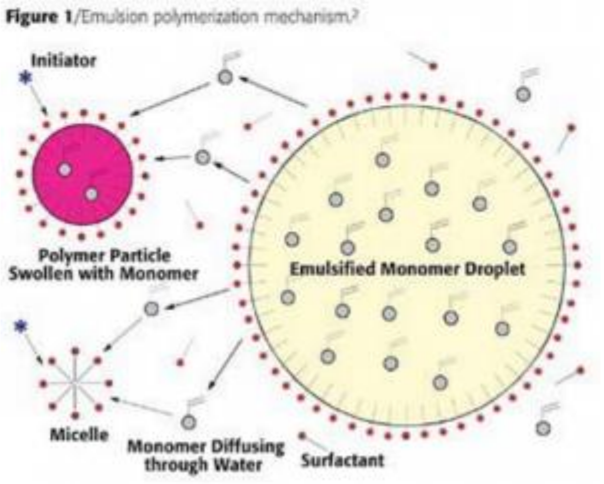


Fig. 11.1

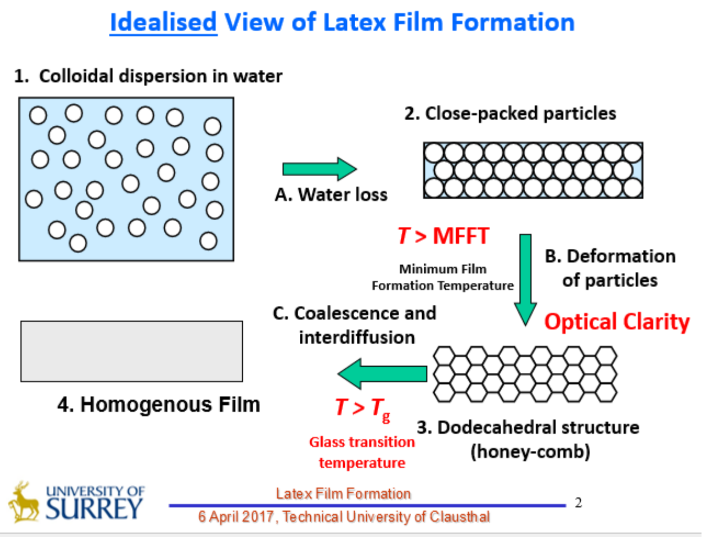
For emulsion polymerization

knowledge.ulprospector.com/4911/pc-fundamentals-emulsion-polymerization/

During Phase I ("water evaporation"), the latex can still flow. This is widely regarded as a problem because noses can form ("sagging"). To prevent this, a thickener is added. The thickener flocculates the dispersion (section 10.10). It is favorable to start film formation with a dispersion, which has a high solids content, already, and just barely flows. This shortens the time during which sagging can occur. Also, the thermal energy required for drying is reduced. (Cost for transport, is reduced as well.) A high solids content is achieved with wide particle size distributions, because the small particles then fit into the gaps (the interstices) between the large particles. Technical dispersions can have a solids content of 50% and higher.

Drying can be accelerated by blowing hot air over the surface. On fast-running belts, adhesive tapes or films are dried within a few seconds. Under ambient conditions, the evaporation rate of water is about 1 mm/day. There is a certain fundamental limit to the drying speed, produced by “laminar boundary layer”. Such a layer (also “unstirred layer”) exists at any interface exposed to flow. Transport through this boundary layer is diffusive, only. This layer can be made thinner by vigorous convection, but it cannot be avoided altogether. (A similar layer in electrochemistry is called the “Nernst layer”. There it is several 100 µm thick). The thickness of the boundary layer and the diffusivity of the water vapor in the boundary layer result in a “mass transfer coefficient”, which limits transport. Similar mass transfer coefficients also exists at all other phase boundaries.

During Phase I, the "coffee-stain effect" or "coffee-ring effect" may occur. In this case, the material flows outward. A bulge is formed at the edge of the film. There are two separate driving mechanisms for the coffee-stain effect. First, the evaporation rate is increased at the edge. Second, the meniscus is anchored at the edge because the solidification starts there. Without this kind of pinning, the droplet would contract laterally. Because of the pinning, material flows outward – and remains there.



Idealised View of Latex Film Formation

1. Colloidal dispersion in water

A. Water loss

2. Close-packed particles

$T >$ MFFT

Minimum Film Formation Temperature

B. Deformation of particles

Optical Clarity

3. Dodecahedral structure (honey-comb)

C. Coalescence and interdiffusion

$T > T_g$

Glass transition temperature

4. Homogenous Film

Latex Film Formation

6 April 2017, Technical University of Clausthal

2

Fig. 11.2
The different phases of film formation from latex dispersions
Kindly provided by Prof. Joe Keddie, University of Surrey

During phase I, there is a concentration gradient of particles along the vertical. This gradient is small if evaporation is slow. Diffusion then efficiently equilibrates the concentration differences. If evaporation is fast, this gradient has profound consequences. Slow or fast evaporation are captured by the "Peclet number",[69] which is defined as

Eq. 11.2.1

$$Pe \approx \frac{\tau_{diff}}{\tau_{dry}} = \frac{H^2 / D}{H / E} = \frac{H}{DE}$$

H is the film thickness, *D* is the diffusivity of the polymer particles. and *E* (in units of m/s) is the evaporation rate. If the solids content is small, one can use the Stokes-Einstein diffusivity for *D* ($D_{SE} = k_B T/(6\pi\eta a)$). Otherwise, one must use the cooperative diffusion coefficient. The latter is defined via Fick's law ($J = -D_{coop}\nabla c$). When repulsive interactions dominate, the cooperative

---

[69] More generally, the Peclet number quantifies the relative importance of diffusion in a transport process. If $Pe << 1$, diffusion dominates the picture.

diffusion coefficient is larger than the Stokes-Einstein diffusivity. Mass transport is more efficient than predicted by Stokes-Einstein because the particles "push each other forward" (hence the name "cooperative").

When the Peclet number is high (when the film dries quickly so that diffusion is ineffective), a skin often forms at the top of the film. This is a problem because this skin may produce cracks or wrinkles, which later turn into film defects. The skin also slows down further drying. Accumulation of particles at the top of the film will only lead to a skin if the particles deform and coalesce there. This will happen for soft particles. The extent of skin formation therefore depends on the glass temperature of the polymer.

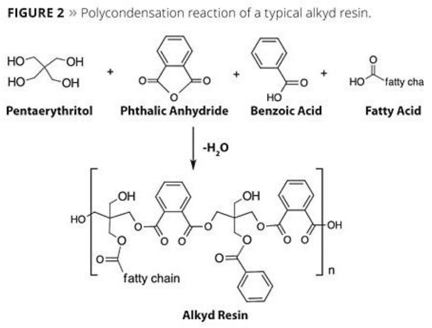


Fig. 11.3
Alkyd resins are the modern technical version of linseed oil, which used to be widely used by painters as a binder. *www.pcimag.com/articles/99678-development-of-low-color-alkyd-resins-with-high-content-of-biobased-succinic-acid*

The Peclet number also plays a role in the process of "auto-stratification" (also "self-stratification"). It would be technically attractive if films with graded properties (functionally graded materials) were to form spontaneously during the drying process. For example, one might want good adhesion on the substrate side and good scratch resistance on the air side. Currently, such coatings are created as multi-layer systems. Coatings in the automotive industry consist of three layers. First, an electrocoat is applied (section 10.4). This layer is optimized for adhesion and corrosion protection. Electrocoating ensures that the paint is deposited everywhere. This is a pH-induced precipitation. The colloids are designed to not be pH stable. It must be a cathodic (rather than an anodic process) because anodic processes tend to dissolve the metal. The pH-unstable colloids are therefore stabilized with amines (instead of carboxylic acids, which would be cheaper). Above this layer is the "base-coat". This layer contains a large portion of the color pigments. At the top is "top-coat", optimized for scratch resistance, often composed of polyurethane.

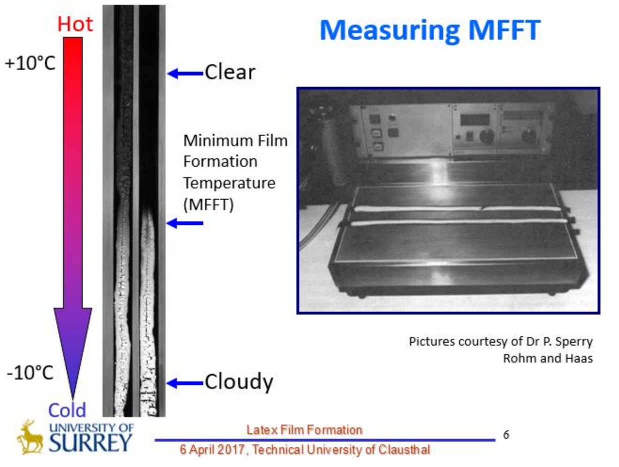


Fig. 11.4
For determining the minimum film forming temperature (MFT). *Kindly provided by Prof. Joe Keddie, University of Surrey.*

The concentration gradients caused by slow diffusion may be frozen if film solidifies early enough. This is the desired “self-stratification”. In order to achieve self-stratification, the Peclet number must be about 1. The particles then initially accumulate at the surface of the film. If one type of particle diffuses faster than the other particle type (e.g. because some particles are smaller than others), the concentration of the small particles equilibrates, while that of the large particles does not. When diffusion arrests, the large particles are enriched at the surface. Following the argument outlined above, the large particles should be enriched at the top. It turns out that sometimes the small particles are enriched at the top. Reasons are not discussed here. Multi-component diffusion can be complicated. In principle, one may also think of a mechanism based on small molecules dissolved in water (in the “serum”). However, small molecules diffuse so quickly that the film would have to a very thick in order to freeze-in the concentration gradients. In other words, H in the numerator of Eq. 11.2.1 must be large to compensate for the large D in the denominator. Such thick films tend to be impractical.

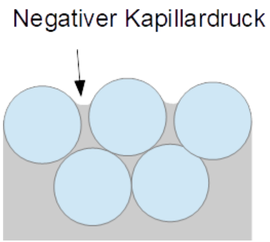


Fig. 11.5
The flow of water may be driven by negative capillary pressure, which occurs at the water/air interface.

We come to phase II of film formation, which is particle deformation. First, particles only deform if they are sufficiently soft. The temperature must not be far below the glass transition temperature. In the context of coatings, this leads to the "film formation dilemma". As a rule, coatings should be mechanically robust. A glass transition temperature above room temperature is desired. But this very requirement stands in the way of film formation, because hard spheres cannot deform. There is a "minimum film formation temperature" (MFT, or MFFT for minimum *f*ilm *f*ormation *t*emperature), which can be determined with a "film drawing bench" (Fig. 11.4). The film drawing bench essentially is a substrate with a temperature gradient.

One of the ways out of the film formation dilemma is to add plasticizers ("coalescing aids") to the dispersion, which evaporate more slowly than the water. These are small molecules. The coalescing aids produce an odor well known from many water-based paints. In the old times, the organic solvents would also smell, but these have largely been eliminated.[70] Actually, the coalescing aids can be about as problematic from an environmental point of view as the organic solvents. Latex dispersions are often portrayed as part of "green chemistry." This is not always the full picture.

Box 11.1

On the study polymer interdiffusion with non-radiative energy transfer between donor and acceptor molecules.

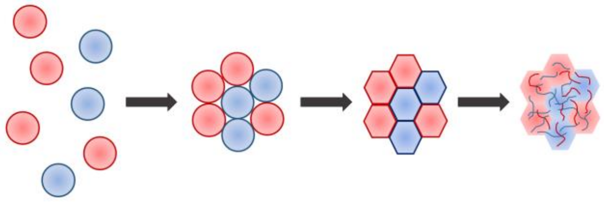


Fig. 11.1

About FRET and interdiffusion

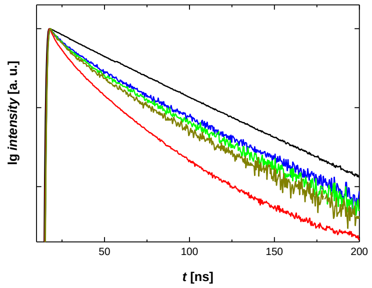


Fig. 11.6
Fluorescence decay curves.

Half of the latex particles are labeled with donor molecules and half with acceptor molecules. The donor dyes fluoresce. The acceptor dyes can fluoresce, but do not always do so. When donor and acceptor dyes are in close proximity, "Förster Resonant Energy Transfer" (FRET) occurs. This is a near-field electromagnetic effect. Electromagnetism is more complicated than what we are used to, when the distance between two antennae is comparable to their size. Then the "near-field effects" come into play. One of these effects is the energy transfer between two dyes. (Magnetic near-field effects are exploited in “near-field communication”, NFC.)

FRET has a characteristic signature in the decay curves. The decay curves are histograms of the time delays between the exciting lamp pulse and the arrival of the fluorescence photon. In the absence of energy transfer, this histogram is close to an exponential. In lin-log representation, it is a straight line. In the presence of FRET, the decay curve bends upward.

---

[70] Do not confuse a solvent with the continuous phase of a dispersion. The polymer in a latex dispersion is insoluble in water.

The second way to cure a soft coating is chemical crosslinking There are several ways to achieve this:

- Radical crosslinking, thermally initiated
- Radical crosslinking, UV-initiated
- Oxygen-induced crosslinking (alkyd resins, Fig. 11.3)
- pH-induced crosslinking, where the pH decreases during drying because ammonia evaporates.

If the particles are *not* sufficiently deformable, two separate problems occur: First, the interstitial phase can fill with air. The film then becomes white. Second, films may crack. Cracks, on the one hand, and a white film, on the other, are seen in Fig. 11.4 as separate phenomena.

Note. Until the interstices fill with air, the films usually are clear. This is true even for small glass spheres (often made by Stöber synthesis, Sect. 10.1). Nanoscale capillary forces suffice to even deform small glass spheres.

There are three different mechanisms that can drive particle deformation. These are negative capillary pressure, which occurs at the water/air interface (Fig. 11.5). Particle deformation can further be driven by the energy of the water-polymer interface ("wet sintering") or by the energy of the air-polymer interface ("dry sintering"). These mechanisms play a role in the classification of film formation scenarios according to Routh and Russel (Fig. 11.8). The kinetics of the sintering process is discussed in **Fehler! Verweisquelle konnte nicht gefunden werden.**.

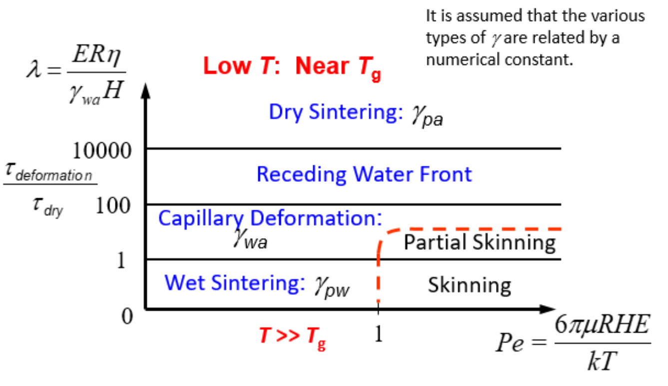
This brings us to phase III of film formation: interdiffusion. When a film has become clear, film formation is not yet complete. The film does not yet have mechanical strength. To create mechanical strength, the polymer chains must be given time to diffuse across particle boundaries and form a transient entanglement network. This process is called "polymer welding" in another context. This process was first studied with neutron scattering. Less demanding instrumentally is a fluorescence method, based on nonradiative energy transfer between donor and acceptor dyes (**Fehler! Verweisquelle konnte nicht gefunden werden.**).

11.3 Different scenarios of film formation

Depending on the conditions, the film formation process can take different courses. Routh and Russel have summarized these scenarios diagrammatically in a "map". They have chosen two main parameters: the Peclet number (Eq. 11.2.1) and a time for particle deformation normalized to the drying time. The various surface tensions are left out as parameters in this representation. The phenomenon of laterally propagating drying fronts is also largely left out. The Routh-Russel model is summarized in Fig. 11.8.

Theory: Deformation mechanism is a function of the dimensionless parameters, λ and *Pe*

It is assumed that the various types of γ are related by a numerical constant.



A.F. Routh & W.B. Russel, *Langmuir*, **15** (1999) 7762-73.

UNIVERSITY OF SURREY

Latex Film Formation

6 April 2017, Technical University of Claustal

2

Fig. 11.7
Different scenarios of film formation according to the Routh-Russel model.
Kindly provided by Prof. Joe Keddie, University of Surrey.