

Physics of Soft Matter

lecture notes

P. Ziherl

*Faculty of Mathematics and Physics, University of Ljubljana
Jožef Stefan Institute*

May 29, 2014

Contents

1	Introduction	7
1.1	Interparticle forces in soft matter*	9
1.2	Viscoelastic response*	11
1.3	Generalized susceptibility	15
2	Simple liquids	19
2.1	Densities and distribution functions*	23
2.2	YBG hierarchy	24
2.3	Energy, pressure, and compressibility	27
2.4	Distribution function theories	29
2.5	Virial expansion	32
2.6	Hard-sphere equation of state*	33
2.7	Perturbation theories	36
3	Liquid crystals	39
3.1	Onsager theory	41
3.2	Nematic elasticity*	45
3.3	Defects*	51
3.4	Nematic order parameter*	53
3.5	Landau-de Gennes theory*	55
3.6	Smectic elasticity	57
4	Polymers	65
4.1	Single polymer chain*	66
4.2	Coil-globule transition	73
4.3	Polymer solutions	75
4.4	Dynamical models	78
4.5	Gels*	83

5	Colloids	89
5.1	Brownian motion*	91
5.2	Interactions between colloids*	93
5.2.1	Casimir interaction	94
5.3	Derjaguin-Landau-Verwey-Overbeek theory*	102
5.4	Phase diagram of hard-sphere colloids	103
6	Amphiphiles	107
6.1	Amphiphile aggregation*	109
6.2	Membrane elasticity	115
6.3	Vesicles*	123
	Homework problems	125
	Bibliography	127

Foreword

These notes were prepared for the one-semester course in theoretical physics of soft condensed matter physics for master students at the Department of Physics, University of Ljubljana. The course consists of 30 hours of lectures and 15 hours of tutorials/seminar. The aim of the course is to provide a broad review the phenomena and the concepts characteristic of soft matter. Each chapter contains a selection of the ideas associated with a given topic — a more comprehensive and in-depth discussion would not fit in the curriculum, and many important questions had to be skipped. The course builds on the standard equilibrium thermodynamics and statistical mechanics as well as on classical elasticity, and the student is expected to have mastered advanced undergraduate courses in these fields.

The notes rely on the material presented in the textbooks listed at the end. I tried to stick to the notation used in these textbooks as much as possible. Chapter 1 (1 week) is based on Ref. [1], chapters 1 and 2 and on Ref. [2], chapter 5; chapter 2 (2 weeks) on Ref. [3], chapters 1-6; chapter 3 (3 weeks) on Ref. [1], chapter 7, on Ref. [15], chapters 2, 3, and 4, on Ref. [16], chapters 6 and 9 [16]; chapter 4 (3 weeks) on Ref. [1], chapter 5, on Ref. [2], chapters 2 and 3, and on Ref. [17], chapter 2; chapter 5 (2 weeks) on Ref. [1], chapter 4; chapter 6 (2 weeks) on Ref. [1], chapter 9 and on Refs. [25] and [26].

The lecture notes are intended as a map of the topics covered and it is advisable to consult the original textbooks as much as possible. The bibliography contains some uncited references, which may serve as supplementary resources. Also included is a list of homework problems.

I will appreciate any comments and suggestions as well as corrections of errors of any kind. The basic sections and subsections marked by asterisks. I thank G. Kahl (Technische Universität Wien) for many helpful suggestions, and T. Dobravec, D. Grošelj, A. Horvat, B. Kavčič, J. Lapajne, J. Mur, G. Posnjak, and K. Vozel for corrections.

Primož Ziherl

Chapter 1

Introduction

Soft matter is a broad term used to refer to materials that are in between the atomic or molecular crystals on one hand and gasses and simple liquids on the other. Soft materials typically contain more than one component (e.g., colloids are micrometer-size particles dispersed in solvent) and they are characterized by mechanical response much softer and much slower than ordinary solids. This gives rise to behavior not encountered in, e.g., crystals.

The distinctive features of soft matter are [1]

- **Lengthscale** Soft materials are based on macro- or supramolecular entities such as polymers, and colloids. Their size is much larger than the size of a single atom and usually smaller than $1 \mu\text{m}$. This means that the details of the interatomic interaction are not very important, coarse-grained interactions are more convenient.
- **Universality** Many aspects of interactions and structures in soft matter does not depend on the specific chemistry but on more general properties. For example, the radius of gyration of polymers does not depend on the chemistry of the monomers but only on the degree of polymerization. It is the linear architecture of polymers that matters. The same applies to membranes. Irrespective of the type of amphiphiles, the elasticity of the bilayer membrane is of the same type in all amphiphiles.
- **Fluctuations** Typical energy scales in soft matter are comparable to thermal energy, which means that the objects are subject to thermal fluctuations. The polymer chain invariably fluctuates in time. They also execute Brownian motion — e.g., colloids are much larger than

atoms but still too small to sediment. As a result, they move in space in a random fashion due to collisions with the molecules of the solvent.

- **Rheology** Many soft-matter fluids are marked by a viscosity much larger than that of simple liquids. In addition, they show viscoelastic behavior which is directly related to the size of particles dispersed in the solvent. This gives rise to non-Newtonian behavior.
- **Polydispersity** Most soft-matter systems are not monodisperse. In polymers, there is always a distribution of degrees of polymerization. In colloidal dispersions, not all particles are the same size, which affects its propensity to crystallize. In this respect, colloidal crystals are very different from atomic crystals. Polydispersity is an inherent property of soft matter.
- **Scaling** Many properties of polymers depend mostly on the number of monomers N , i.e., the degree of polymerization. One is typically interested in how do these properties scale with N rather than on the exact numerical predictions.
- **Self-assembly or self-organization** Crucial in soft matter is the presence of the solvent, and the interaction with the solvent as well as interactions mediated by the solvent often drive the particles to self-assemble in various ways. Amphiphilic molecules form bilayer membranes or micelles such that the hydrophobic tails are not in contact with water, and these membranes may assume a range of shapes themselves. Depending on concentration, diblock copolymer solutions may form micelles, cylinders, lamellae, gyroids...

The various types of soft-matter materials are:

- **Liquid crystals** Formed of rod-like or plate-like molecules, liquid crystals are viscous liquid characterized by orientational molecular order. Their optical properties are similar to those seen in crystals (especially birefringence). There exist several types of LC phases, the simplest type are nematics.
- **Polymers** are macromolecules of typically linear architecture. In solution, they may form coils or globules. They may be adsorbed on surfaces etc. There also exist more complex polymers — star polymers, dendrimers etc.

- **Colloids** are small μm -size particles dispersed in the continuous medium. Prototype colloids are sols = dispersions of solid particles in a liquid. Colloids can form crystals, colloidal glasses etc.
- **Amphiphile self-assemblies** In water, amphiphilic molecules arrange so that the hydrophilic head faces water and the hydrophobic tail points away from water. This can give rise to bilayer membranes and many other complex structures.
- **Granular materials** are similar to colloids but the particles are bigger so that the gravitational potential energy does matter. Thermal motion is frozen, an interesting phenomenon are avalanches.
- **Foams** are colloids with gas bubbles dispersed in liquid. The border between bubbles are called Plateau borders and they satisfy the Plateau rules.

As many of these materials consist of the solvent, some insight into the theory of simple liquids is needed too.

Many of these types of order are seen in everyday life either in natural or in man-made materials: honey, toothpaste, whipped egg white, mayonnaise, rubber. . .

1.1 Interparticle forces in soft matter*

The very existence of condensed phases (liquids and solids) witnesses to the existence of attractive forces between the building blocks of matter. A liquid film minimizes its surface area at constant volume so as to minimize the fraction of exposed molecules at the surface which are bound less strongly than the molecules in the bulk. On the other hand, condensed matter resists compression so that at small separations, there exist repulsive intermolecular forces. A typical force profile is shown in Fig. 1.1.

The microscopic origin of the short-range repulsion is the electron-electron interaction and the Pauli exclusion principle. This is why the atoms cannot penetrate one another, and this also applies to molecules (albeit in a more complicated way because their conformation may change due to interaction with neighbors). An often-used approximation for the short-range repulsion is the **hard-core potential**.

There exist several types of long-range interactions in soft matter listed below.

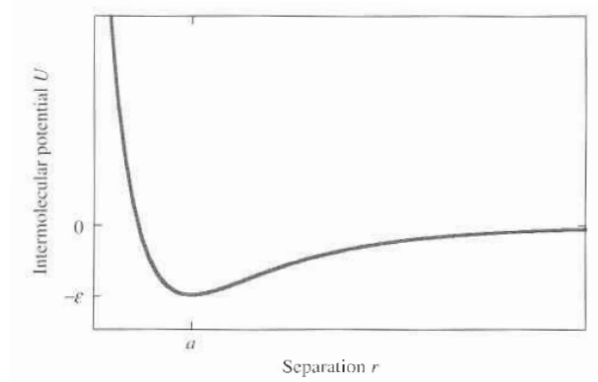


Figure 1.1: Qualitative dependence of the intermolecular force on separation [1].

van der Waals force*

These forces are due to the interaction of an instantaneous rather than permanent dipole moment in a given neutral atom and an induced dipole moment in another neutral atom, which necessarily points in a direction such that the interaction is attractive. The van der Waals force between atoms is given by

$$U_{vdW} = \frac{A}{r^6}. \quad (1.1)$$

The r^{-6} dependence is characteristic of this force. Just like atoms, microscopic bodies also interact with one another with the van der Waals interaction.

The van der Waals interaction is often referred to as the dispersion interaction. For molecules in contact, its magnitude is about 10^{-20} J i.e. similar to $k_B T$.

Electrostatic force*

Most if not all soft-matter systems are solutions of some kind, and in such an environment the basic Coulombic interaction

$$U_C = \frac{q_1 q_2}{4\pi\epsilon_0 r} \quad (1.2)$$

is more complicated than in simple bodies. The reason for this is that we are usually interested in the interaction between macromolecules or colloids suspended in the solution with dissolved counterions. Unlike the interaction

between bare charges, the interaction between these bodies is not pairwise additive because it depends on the distribution of the counterions which in turn depends on the location of the macromolecules. In addition, the counterions screen the interaction so that it decays exponentially with distance.

An order of magnitude estimate of the electrostatic interaction between ions in a crystal: $q_1 = q_2 = e_0$, $r = 0.1$ nm gives 2.3×10^{-18} J. This is much more than $k_B T = 4 \times 10^{-21}$ J. But in solution the electrostatic interaction is much weaker.

Covalent bonds*

These bonds bind atoms into molecules, and are of the same order of magnitude as the bare ion-ion electrostatic interaction, i.e., $\sim 10^{-18}$ J. They are much stronger than $k_B T$ and essentially unaffected by thermal motion. The same applies to metallic bonds.

Hydrogen bonds*

Hydrogen bonds occur between the largely positively hydrogen atoms in molecules (which are reduced to protons as the electron resides at other bigger atoms) and electronegative atoms in neighboring molecules. An typical example is the bond between *H* and *O* atoms in water. Typical energies are up to 6×10^{-20} J, i.e., up to $10k_B T$.

Hydrophobic interaction*

A foreign molecule or inclusion in water changes the local structure of water, and the interaction of the perturbed regions around two nearby inclusions leads to attraction between them. The detailed mechanism of hydrophobic interaction is complicated and still not completely clear. The magnitude of this interaction is $\sim 10^{-20}$ J.

1.2 Viscoelastic response*

One of the hallmark features of soft matter is that the distinction between solids and liquids is not as clear as in simple solids and simple liquids. To understand these differences, we need to look into the solids and liquids in more detail. Specifically, we must understand the workings of the response to shear stress.

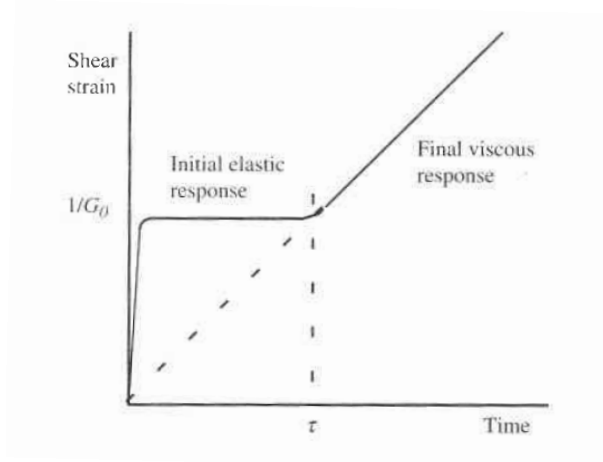


Figure 1.2: Displacement upon an application of a step-like shear stress: $\eta = G(0)\tau$ [1].

In simple condensed matter — crystals and simple liquids — the distinction between solids and liquids is very clear. In **Hookean solids** an applied shear stress produces a shear strain, which is proportional to stress and inversely proportional to Young modulus. In **Newtonian liquids** an applied shear stress produces flow proportional to stress and inversely proportional to viscosity.

These two types of behavior are the limiting cases of a more general response of materials to shear stress known as the viscoelasticity. For reasons described below, it is characteristic of materials formed by large supramolecular entities such as polymers. In viscoelastic materials the response qualitatively depends on the frequency or the time scale. At short times/high frequencies, the material behaves as a solid whereas at long times/low frequencies it flows as a liquid. The typical strain-time dependence is shown in Fig. 1.2.

Apart from viscoelastic behavior, the effective viscosity of complex fluids depends on strain rate (i.e., the shear velocity v/l). In Newtonian fluids it is constant, whereas in shear-thinning fluids it decreases with strain rate. Examples of shear-thinning fluids are paints. When applied with a brush, the rate is high and the paint readily flows. Once on the wall, the rate is small and the viscosity increases. On the other hand, shear-thickening fluids become more resistant to flow at high strain rates. This is characteristic of pastes of a large volume fraction of particles.

Microscopic interpretation*

The elastic modulus and the viscosity of materials can be qualitatively understood in terms of simple models. Let us consider uniaxial deformation along z axis. In equilibrium, the particles are located in the minima of the potential created by their neighbors and their energy can be expanded around the equilibrium particle-particle distance a :

$$U(r) = U(a) + \frac{1}{2} \left. \frac{d^2U}{dr^2} \right|_{r=a} (r - a)^2 + \dots \quad (1.3)$$

The force on a particle is given by

$$F = -\frac{dU}{dr} = -\left. \frac{d^2U}{dr^2} \right|_{r=a} (r - a). \quad (1.4)$$

If we assume that the particles are arranged in a cubic lattice, the stress is given by force per spring divided by area per cell a^2 . So the stress is

$$p_{zz} = \frac{F}{a^2} = -\left. \frac{d^2U}{dr^2} \right|_{r=a} \frac{(r - a)}{a^2} \quad (1.5)$$

whereas the strain is

$$u_{zz} = \frac{r - a}{a}. \quad (1.6)$$

The Young modulus is thus

$$E = \frac{p_{zz}}{u_{zz}} = \frac{1}{a} \left. \frac{d^2U}{dr^2} \right|_{r=a}. \quad (1.7)$$

To estimate the order of magnitude, we further assume that the potential U can be written as

$$U(r) = \epsilon f(r/a), \quad (1.8)$$

where ϵ is the depth of the potential at $r = a$ and f is a function which has a minimum at $r = a$, diverges at $r \rightarrow 0$ and goes to 0 as $r \rightarrow \infty$. In this case, $d^2U/dr^2|_{r=a} \approx \epsilon/a^2$ if we assume that $f''(1) = 1$ and so

$$E \approx \frac{\epsilon}{a^3}. \quad (1.9)$$

Thus the Young modulus is given by the energy density. In soft matter, the interparticle bonds are soft and that is why the elastic modulus is small. More specifically, in an elemental solid $\epsilon \sim 10^{-18}$ J (covalent bonds) and

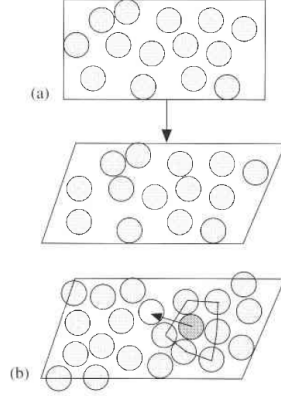


Figure 1.3: Cages in fluid state: Upon shear deformation, the local cage around a given particle is deformed and particle rattles within it before escaping to relax the stress [1].

$a \sim 10^{-10}$ m. This gives $E \sim 10^{12}$ N/m². In a colloidal crystal $\epsilon \sim 10^{-20}$ J and $a \sim 10^{-6}$ m, which gives $E \sim 10^{-2}$ N/m².

In a fluid state, applied stress increases the local energy density and until the particles rearrange, the material sustains stress, i.e., it behaves as a solid. In simple liquids, the rearrangement time is very short and the solid-like behavior is not very prominent. In polymers and colloidal suspension, the rearrangement time is much longer and the short-time elastic response is readily seen.

Upon application of stress, particles caught in a cage formed by neighbors try to jump out of the cage. The probability for escape can be estimated as follows. The particles bounce between the neighbors with a certain frequency ν and the potential barrier for escape is of the same order but smaller than ϵ (the depth of the potential minimum which is the heat of vaporization per particle). The probability for escape per unit time (i.e., the inverse relaxation time) is thus

$$\frac{1}{\tau} \sim \nu \exp\left(-\frac{\epsilon}{k_B T}\right). \quad (1.10)$$

Now the heat of vaporization per particle in small molecules is smaller than in polymers, and ϵ enters the above equation in the exponential. This is why the relaxation time in simple liquids is between 10^{-12} and 10^{-10} s whereas in polymers it can be up to ms.

Upon application of shear stress F/A

$$\frac{\Delta l}{l} = \frac{1}{G} \frac{F}{A}, \quad (1.11)$$

where $\Delta l/l$ is the shear strain and G is the shear modulus.

The displacement of a viscoelastic material to step-function shear stress consists of two parts: i) the initial step function due to the elastic response and ii) the linear viscous response which kicks in after the relaxation time τ . In the initial elastic regime the displacement does not depend on time whereas in the viscous regime it increases linearly with time (Fig. 1.2). The magnitude of the short-time displacement is inversely proportional to the static shear modulus $G(0)$ (and proportional to α')

$$\frac{\Delta l}{l} = \frac{1}{G(0)} \frac{F}{A}, \quad (1.12)$$

where $\Delta l/l$ is the shear strain and F/A is shear stress. In the viscous regime

$$\frac{v}{l} = \frac{1}{\eta} \frac{F}{A}, \quad (1.13)$$

where η is the viscosity. By comparing the two regimes we find that

$$\eta = G(0)\tau. \quad (1.14)$$

Equation (1.10) suggests that

$$\eta = \frac{G(0)}{\nu} \exp\left(\frac{\epsilon}{k_B T}\right) \quad (1.15)$$

This is called the Arrhenius behavior.

1.3 Generalized susceptibility

Let us describe the elastic and the viscous response of the system to the generalized force ψ in terms of the generalized coordinate (deformation or displacement) x [2]. The system is characterized by the primary response function $\mu(t)$ which describes the displacement due to an instantaneous force $\psi(t) = \psi_0 \delta(t)$:

$$x(t) = \mu(t)\psi_0. \quad (1.16)$$

The typical response function is shown in Fig. 1.4. The delta function describes a solid where the displacement is gone as soon as the force ceases to

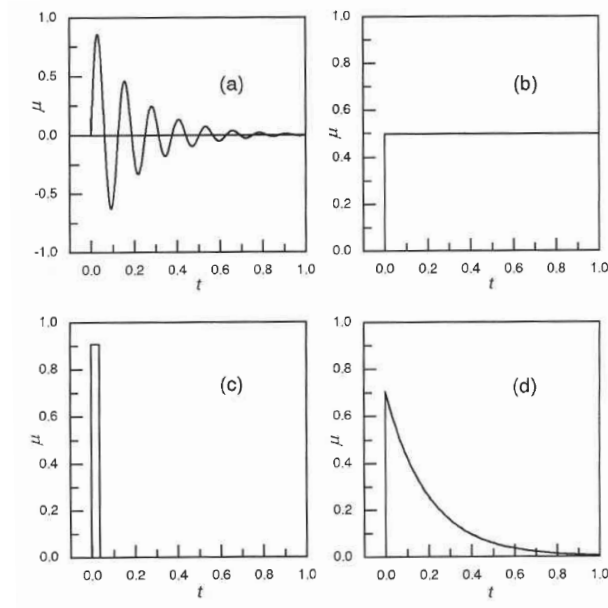


Figure 1.4: Primary response function of a damped harmonic oscillator (a), a liquid (b), a solid (c), and a simple relaxatory system (d) [2].

act. The Heaviside function describes a viscous body where the displacement persists even after the force ceases to act. Real liquids are described by a rounded step function, say $1 - \exp(-t/\tau)$ where τ is the relaxation time. In relaxatory systems the displacement decays with time.

The response of the system to arbitrary force is then given by

$$x(t) = \int_{-\infty}^t \mu(t-t')\psi(t')dt'. \quad (1.17)$$

This formula follows from

- **Causality:** The response depends on forces in the past so that the upper limit of the integral is t .
- **Superposition:** Any time-dependent force can be written as a sequence of δ -functions of suitable amplitudes, and the total response depends on all of them.

Consider the response to an oscillatory force

$$\psi(t) = \psi_0 \exp(i\omega t). \quad (1.18)$$

The displacement oscillates with the same frequency but is phase-shifted

$$x(t) = x_0 \exp(i\omega t) \exp(-i\phi). \quad (1.19)$$

We substitute these formulas in Eq. (1.17):

$$x_0 \exp(i\omega t) \exp(-i\phi) = \int_{-\infty}^t \mu(t-t') \psi_0 \exp(i\omega t') dt'. \quad (1.20)$$

Now we define the dynamic susceptibility by

$$\alpha^* = \frac{\text{displacement}}{\text{force}} = \frac{x_0 \exp(-i\phi)}{\psi_0} = \int_{-\infty}^t \mu(t-t') \exp(i\omega(t'-t)) dt'. \quad (1.21)$$

By replacing $t-t' = t''$ we have $\alpha^* = \int_0^\infty \mu(t'') \exp(-i\omega t'') dt''$. Since the primary response function vanishes at $t < 0$, $\mu(t < 0) = 0$, the lower limit of the integral can be replaced by $-\infty$. We find that

$$\alpha^* = \int_{-\infty}^\infty \mu(t'') \exp(-i\omega t'') dt''. \quad (1.22)$$

Thus the dynamic susceptibility is the Fourier transform of the primary response function. $\alpha^* = \alpha' - i\alpha''$ is a complex function and its real and imaginary parts α' and α'' are not independent as they are related to the same primary response function. They satisfy the Kramers-Kronig relations

$$\alpha'(\omega_0) = \frac{1}{\pi} P \int_{-\infty}^\infty \frac{\alpha''}{\omega - \omega_0} d\omega, \quad (1.23)$$

$$\alpha''(\omega_0) = -\frac{1}{\pi} P \int_{-\infty}^\infty \frac{\alpha'}{\omega - \omega_0} d\omega. \quad (1.24)$$

Let us look at two typical examples:

- **Liquid:** $\mu(t) = \Theta(t)$, $\alpha^* = \pi\delta(\omega) - i/\omega$. Imaginary susceptibility is nonzero at all frequencies, losses prominent. Real susceptibility is nonzero only at $\omega = 0$ which corresponds to static force.
- **Solid:** $\mu(t) = \delta(t)$, $\alpha^* = 1$. Imaginary susceptibility vanishes, no losses.

Now we calculate the work done by an oscillatory force $\psi(t) = \psi_0 \cos(\omega t)$. The displacement reads $x(t) = \alpha^* \psi(t) = (\alpha' - i\alpha'') \psi(t)$, its real part being

$$x(t) = \alpha' \psi_0 \cos \omega t + \alpha'' \psi_0 \sin \omega t. \quad (1.25)$$

The power reads

$$\frac{dW}{dt} = \psi \frac{dx}{dt} = \psi_0 \cos \omega t (-\omega \alpha' \psi_0 \sin \omega t + \omega \alpha'' \psi_0 \cos \omega t) \quad (1.26)$$

$$= -\frac{\psi_0^2}{2} \omega \alpha' \sin 2\omega t + \psi_0^2 \omega \alpha'' \cos^2 \omega t. \quad (1.27)$$

The first term oscillates with time and its average over one cycle is 0 — obviously it represents the energy which is stored in the system during the first half-cycle and retrieved during the second half-cycle. Thus α' is related to the elastic response of the system.

The second term is qualitatively different, its time average being

$$\overline{\frac{dW}{dt}} = \frac{\psi_0^2}{2} \omega \alpha'' > 0. \quad (1.28)$$

In an isothermal system, the internal energy is constant and from $dU = dW + dQ$ it follows that the work expended on the system $dW > 0$ and the heat released into the thermostat $dQ > 0$ are proportional to α'' .

Chapter 2

Simple liquids

By themselves, simple liquids do not belong to soft matter. Yet many soft materials either are liquids or are based on liquids, and thus some understanding of the liquid state is needed. In addition, the density of liquids is very similar to the density of the solids, so that the molecules interact strongly and the partition function cannot be evaluated exactly. Instead a series expansion of the density and the distribution function are used — and typically truncated at the pair level. The pair distribution at a given pair interaction is determined by the first term in YBG hierarchy, leading to an integrodifferential equation. Also used to determine $g(r)$ is the Ornstein-Zernike equation, and the equation of state and the thermodynamic potentials are computed from the energy, the pressure, and the compressibility equations. This usually leads to thermodynamic inconsistency due to the various approximations made.

Here we will review some results presented in Hansen and McDonald's textbook [3].

Liouville equation*

The liquid consists of N classical spinless particles and the Hamiltonian reads

$$H_N = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + V_N(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) + \sum_{i=1}^N V_{ext}(\mathbf{r}_i). \quad (2.1)$$

The first term is the kinetic energy, the second one describes the interparticle interactions, and the last one is the external potential.

The statistical-mechanical system is described by the phase-space probability density $f^{(N)}(\mathbf{r}^N, \mathbf{p}^N; t)$. The time derivative of f must vanish as the

probability density in phase space is conserved. f depends on time both directly and indirectly so that

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_{i=1}^N \left(\frac{\partial f}{\partial \mathbf{r}_i} \cdot \frac{d\mathbf{r}_i}{dt} + \frac{\partial f}{\partial \mathbf{p}_i} \cdot \frac{d\mathbf{p}_i}{dt} \right) = 0. \quad (2.2)$$

This is the Liouville equation. The Hamiltonian dynamics read

$$\frac{d\mathbf{r}_i}{dt} = \frac{\partial H_N}{\partial \mathbf{p}_i} = \frac{1}{m} \mathbf{p}_i. \quad (2.3)$$

and

$$\frac{d\mathbf{p}_i}{dt} = -\frac{\partial H_N}{\partial \mathbf{r}_i} = \mathbf{F}_i \quad (2.4)$$

where \mathbf{F}_i is the force on i th particle; to simplify the discussion, we set the external potential to 0. Upon inserting these results in Eq. (2.2) we find

$$\frac{\partial f}{\partial t} + \sum_{i=1}^N \left(\frac{1}{m} \frac{\partial f}{\partial \mathbf{r}_i} \cdot \mathbf{p}_i + \frac{\partial f}{\partial \mathbf{p}_i} \cdot \mathbf{F}_i \right) = 0. \quad (2.5)$$

In equilibrium, $\partial f / \partial t = 0$.

Partition function and configuration integral*

The canonical partition function is

$$Q_N(V, T) = \frac{1}{N! h^{3N}} \int \exp(-\beta H_N(\mathbf{r}^N, \mathbf{p}^N)) d\mathbf{r}^N d\mathbf{p}^N. \quad (2.6)$$

Here V and T are volume and temperature, respectively, and \mathbf{r}^N and \mathbf{p}^N are vectors of the positions and momenta of all particles. The Helmholtz free energy is given by

$$F = -k_B T \ln Q(V, T) \quad (2.7)$$

and pressure and entropy follow from $dF = -SdT - pdV$:

$$p = - \left(\frac{\partial F}{\partial V} \right)_T \quad (2.8)$$

and

$$S = - \left(\frac{\partial F}{\partial T} \right)_V. \quad (2.9)$$

In terms of the probability density $\exp(-\beta H_N(\mathbf{r}^N, \mathbf{p}^N)) / (N!h^{3N})$, the internal energy U reads

$$U = \frac{1}{N!h^{3N}Q_N(V,T)} \int H_N(\mathbf{r}^N, \mathbf{p}^N) \exp(-\beta H_N(\mathbf{r}^N, \mathbf{p}^N)) \, d\mathbf{r}^N \, d\mathbf{p}^N. \quad (2.10)$$

The kinetic part of the partition function can be integrated readily giving

$$Q_N(V,T) = \frac{1}{N!\Lambda^{3N}} Z_N(V,T), \quad (2.11)$$

where

$$\Lambda = \frac{h}{\sqrt{2\pi m k_B T}} \quad (2.12)$$

is the thermal de Broglie wavelength and

$$Z_N(V,T) = \int \exp(-\beta V_N(\mathbf{r}^N)) \, d\mathbf{r}^N \quad (2.13)$$

is the configuration integral.

In ideal gas, $V_N(\mathbf{r}^N) = 0$ and $Z_N(V,T) = V^N$ so that the ideal free energy is

$$F^{id} = -k_B T \ln \left(\frac{V^N}{N!\Lambda^{3N}} \right) \quad (2.14)$$

$$\approx -k_B T (N \ln V - N \ln N + N - 3N \ln \Lambda) \quad (2.15)$$

$$= N k_B T (\ln \rho + 3 \ln \Lambda - 1) \quad (2.16)$$

where we used the Stirling approximation $\ln N! \approx N \ln N - N$ and $\rho = N/V$. This readily gives the ideal-gas equation of state $pV = N k_B T$.

Ideal and excess quantities*

In a system of interacting particles,

$$Q_N(V,T) = Q_N^{id}(V,T) \frac{Z_N(V,T)}{V^N} \quad (2.17)$$

and the free energy given by Eq. (2.7) naturally splits into the ideal and the excess parts

$$F = F^{id} + F^{ex}, \quad (2.18)$$

where

$$F^{ex} = -k_B T \ln \frac{Z_N(V,T)}{V^N}. \quad (2.19)$$

In a similar way, we find that $U = U^{id} + U^{ex}$ where $U^{id} = 3Nk_B T/2$ and

$$U^{ex} = \frac{1}{Z_N(V, T)} \int V_N(\mathbf{r}^N) \exp(-\beta V_N(\mathbf{r}^N)) \mathbf{d}\mathbf{r}^N = \langle V_N \rangle. \quad (2.20)$$

Grand canonical ensemble

The grand-canonical partition function reads

$$\Xi(V, T) = \sum_{N=1}^{\infty} \frac{\exp(N\beta\mu)}{N! h^{3N}} \int \exp(-\beta H_N(\mathbf{r}^N, \mathbf{p}^N)) \mathbf{d}\mathbf{r}^N \mathbf{d}\mathbf{p}^N \quad (2.21)$$

$$= \sum_{N=1}^{\infty} \exp(N\beta\mu) Q_N(V, T) \quad (2.22)$$

$$= \sum_{N=1}^{\infty} \frac{z^N}{N!} Z_N(V, T), \quad (2.23)$$

where $z = \exp(\beta\mu)/\Lambda^3$ is the activity and μ is the chemical potential.

The chemical potential can be expressed in terms of derivatives of U , F , and G :

$$\mu = \left(\frac{\partial U}{\partial N} \right)_{V, S} = \left(\frac{\partial F}{\partial N} \right)_{V, T} = \left(\frac{\partial G}{\partial N} \right)_{p, T}. \quad (2.24)$$

Like U and F , it can be split into the ideal and the excess part. Using Eq. (2.19), the latter can be computed from

$$\mu^{ex} = \left(\frac{\partial F^{ex}}{\partial N} \right)_{V, T} = F^{ex}(N+1, V, T) - F^{ex}(N, V, T) \quad (2.25)$$

$$= -k_B T \ln \left(\frac{Z_{N+1}(V, T)}{V Z_N(V, T)} \right). \quad (2.26)$$

The potential energy of $N+1$ particles $V_{N+1}(\mathbf{r}^N)$ can be divided into $V_N(\mathbf{r}^N)$ and ϕ

$$V_{N+1}(\mathbf{r}^N) = V_N(\mathbf{r}^N) + \phi, \quad (2.27)$$

where ϕ is the interaction of $(N+1)$ th particle with all others. In a translationally invariant system \mathbf{r}_{N+1} may be taken as the origin of the coordinate system. Integrating over \mathbf{r}_{N+1} gives a factor of V and we have

$$\frac{Z_{N+1}(V, T)}{Z_N(V, T)} = \frac{V \int \exp(-\beta V_N(\mathbf{r}^N)) \exp(-\beta\phi) \mathbf{d}\mathbf{r}^N}{\int \exp(-\beta V_N(\mathbf{r}^N)) \mathbf{d}\mathbf{r}^N} = V \langle \exp(-\beta\phi) \rangle. \quad (2.28)$$

Thus the excess chemical potential

$$\mu^{ex} = -k_B T \ln \langle \exp(-\beta\phi) \rangle \quad (2.29)$$

is proportional to the logarithm of the mean Boltzmann factor of the test particle inserted in a system of N particles.

2.1 Densities and distribution functions*

The n -particle density is defined by

$$\begin{aligned} \rho_N^{(n)}(\mathbf{r}^n) &= \frac{N!}{(N-n)! N! h^{3N} Q_N(V, T)} \\ &\quad \times \int \exp(-\beta H_N(\mathbf{r}^N, \mathbf{p}^N)) d\mathbf{r}^{(N-n)} d\mathbf{p}^N \end{aligned} \quad (2.30)$$

$$= \frac{N!}{(N-n)! Z_N(V, T)} \int \exp(-\beta V_N(\mathbf{r}^N)) d\mathbf{r}^{(N-n)}. \quad (2.31)$$

The normalization is such that

$$\int \rho_N^{(n)}(\mathbf{r}^n) d\mathbf{r}^n = \frac{N!}{(N-n)!}. \quad (2.32)$$

This gives

$$\int \rho_N^{(1)}(\mathbf{r}) d\mathbf{r} = N \quad (2.33)$$

and in a homogeneous system

$$\rho_N^{(1)}(\mathbf{r}) = \frac{N}{V} = \rho. \quad (2.34)$$

In ideal gas, $V_N(\mathbf{r}^N) = 0$ and $Z_N(V, T) = V^N$ so that Eq. (2.31) gives

$$\rho_N^{(n)}(\mathbf{r}^n) = \frac{N!}{(N-n)! V^n} = \frac{N!}{(N-n)! N^n} \rho^n. \quad (2.35)$$

The first factor can be written as

$$\begin{aligned} \frac{N(N-1)\dots(N-n+1)}{N^n} &= \left(1 - \frac{1}{N}\right) \dots \left(1 - \frac{n-1}{N}\right) \\ &\sim 1 + \mathcal{O}\left(\frac{n}{N}\right) \end{aligned} \quad (2.36)$$

so that in ideal gas the pair density

$$\rho_N^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \rho^2 \left(1 - \frac{1}{N}\right). \quad (2.37)$$

In a real gas where the only term in the potential $V_N(\mathbf{r}^N)$ that has not been integrated out is the pair potential $v(\mathbf{r}_1, \mathbf{r}_2) = v(r = |\mathbf{r}_1 - \mathbf{r}_2|)$ so that

$$\rho_N^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \rho^2 \left(1 - \frac{1}{N}\right) \exp(-\beta v(r)). \quad (2.38)$$

The n -particle distribution function is defined by

$$g_N^{(n)}(\mathbf{r}^n) = \frac{\rho_N^{(n)}(\mathbf{r}^n)}{\prod_{i=1}^n \rho_N^{(1)}(\mathbf{r}_i)} \quad (2.39)$$

and in a homogeneous system

$$g_N^{(n)}(\mathbf{r}^n) = \rho^{-n} \rho_N^{(n)}(\mathbf{r}^n). \quad (2.40)$$

The most important distribution function is the pair distribution function $g_N^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$. In an isotropic and homogeneous system, $g_N^{(2)}$ depends only on $r = |\mathbf{r}_1 - \mathbf{r}_2|$ and is called the radial distribution function. From Eq. (2.38) we have

$$g_N^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \exp(-\beta v(r)), \quad (2.41)$$

where we have assumed that $N \gg 1$. For r much larger than the range of interparticle potential, $g(r)$ tends to 1. The typical radial distribution function is shown in Fig. 2.1.

An especially important example is the hard-sphere fluid. The exact radial distribution function can only be obtained numerically and the result is shown in Fig. 2.2. As density is increased, the anticorrelation hole appears after the nearest-neighbor peak and so do the second- and third-nearest neighbor peaks.

2.2 YBG hierarchy

How do the distribution functions depend on the pair potential? This is described by the Liouville equation [Eq. (2.5)]. In brief, the equilibrium probability density $f^{(N)}$ is factorized — the momenta and the kinetic energy

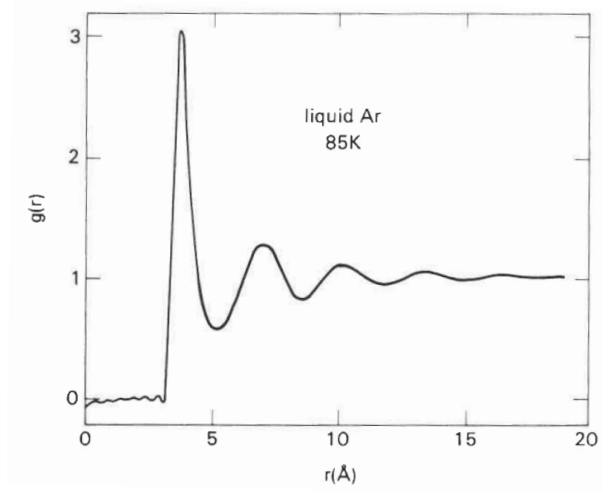


Figure 2.1: Radial distribution function in liquid argon [3].

Figure 2.2: Radial distribution function in hard spheres at various densities. Monte-Carlo results.

gives the standard Maxwell-Boltzmann distribution and $\rho^N(\mathbf{r}^N)$ captures the effect of interactions between the particles:

$$f^{(N)}(\mathbf{r}^N, \mathbf{p}^N) = \rho^{(N)}(\mathbf{r}^N) \left(\frac{\beta}{2\pi m} \right)^{3N/2} \exp \left(-\beta \sum_{i=1}^N \mathbf{p}_i^2 / 2m \right). \quad (2.42)$$

After this is inserted in the equilibrium Liouville equation, we obtain a differential equation for $\rho^{(N)}(\mathbf{r}^N)$ which features the forces on each particle. This force can be expressed in terms of forces of all other particles.

If the total density is expressed in terms of the distribution functions, one obtains the Yvon-Born-Green (YBG) hierarchy of recursive integrodifferential scheme which relate, e.g., $g^{(n)}$ and $g^{(n+1)}$. The derivation is rather involved and we only list the first and the most important member of this hierarchy which for a homogeneous liquid reads

$$\nabla_1 g(r_{12}) = -\beta \nabla_1 v(r_{12}) g(r_{12}) - \beta \rho \int \nabla_1 v(r_{13}) g^{(3)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) d\mathbf{r}_3. \quad (2.43)$$

This is essentially the force equation for the liquid. In the superposition approximation,

$$g^{(3)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = g(r_{12}) g(r_{13}) g(r_{23}) \quad (2.44)$$

and upon dividing Eq. (2.43) by $g(r_{12})$ and rearranging we have

$$\nabla_1 [\ln g(r_{12}) + \beta v(r_{12})] = -\beta \rho \int \nabla_1 v(r_{13}) g(r_{13}) g(r_{23}) d\mathbf{r}_3. \quad (2.45)$$

Finally, we note that in an isotropic fluid $\int \nabla_1 v(r_{13}) g(r_{13}) d\mathbf{r}_3 = 0$ so that the right-hand side can be written as

$$\nabla_1 [\ln g(r_{12}) + \beta v(r_{12})] = -\beta \rho \int \nabla_1 v(r_{13}) g(r_{13}) h(r_{23}) d\mathbf{r}_3, \quad (2.46)$$

where $h(r) = g(r) - 1$ is the pair correlation function. By solving this equation at a given pair potential we obtain the radial distribution function.

Here only the full series of distribution functions provides a complete description. If the series is truncated and we only use the pair distribution function we obtain inconsistent results. A way around this set is to define the pair and the total correlation function. This leads to Ornstein-Zernike equation.

2.3 Energy, pressure, and compressibility

In a system of particles with pairwise interaction $v(\mathbf{r}_{ij})$ the total potential energy is

$$V_N(\mathbf{r}^N) = \frac{1}{2} \sum_{\substack{i,j \\ i \neq j}} v(\mathbf{r}_{ij}). \quad (2.47)$$

By recognizing that the average number of particles separated from the reference particle by r is $n(r)dr = 4\pi r^2 \rho g(r)dr$ and that the potential energy of the reference particle due to particles separated by r is $v(r)n(r)dr$, we find that the excess internal energy reads

$$U^{ex} = 2\pi N \rho \int_0^\infty v(r)g(r)r^2 dr, \quad (2.48)$$

where we took into account that the interaction of a pair of particles should be counted only once. This is the energy equation and can also be obtained in a more formal way [3].

The pressure equation is derived from the Clausius virial function

$$\mathcal{V}(\mathbf{r}^N) = \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{F}_i, \quad (2.49)$$

where \mathbf{F}_i is the total force on particle i at \mathbf{r}_i . The time average of $\mathcal{V}(\mathbf{r}^N)$ is

$$\langle \mathcal{V} \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau d\tau \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{F}_i \quad (2.50)$$

$$= \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau d\tau \sum_{i=1}^N \mathbf{r}_i \cdot m \ddot{\mathbf{r}}_i \quad (2.51)$$

$$= - \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau d\tau \sum_{i=1}^N m |\dot{\mathbf{r}}_i|^2 \quad (2.52)$$

$$= -3Nk_B T, \quad (2.53)$$

where we integrated by parts and finally used the equipartition theorem: The average value of twice the kinetic energy is $3Nk_B T$.

\mathcal{V} can be divided into the external part arising from the outside pressure and the \mathcal{V}_{int} due to the interaction between the particles. The former can be computed by considering a cuboidal volume L^3 exposed to hydrostatic pressure p . The virial function is then the work done by the pressure within

the liquid against the forces $pA = pL^2$ along the $x, y,$ and z axis. The cube sides are displaced by L so that the work done is $-3pAL = -3pV$. By substituting this results in $\mathcal{V} = \mathcal{V}_{int} + \mathcal{V}_{ext} = \mathcal{V}_{int} - 3pV$ into Eq. (2.53) and dividing it by 3 we find

$$pV = Nk_B T + \frac{1}{3}\langle \mathcal{V}_{int} \rangle = Nk_B T - \frac{1}{3}\left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \nabla_i V_N(\mathbf{r}^N) \right\rangle \quad (2.54)$$

or

$$\frac{\beta p}{\rho} = 1 - \frac{\beta}{3N}\left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \nabla_i V_N(\mathbf{r}^N) \right\rangle. \quad (2.55)$$

This is the pressure or virial equation. Since all particles are equivalent, the average contains N identical terms. The number of particles in a shell of volume $4\pi r^2 dr$ at a certain r is $4\pi\rho g(r)r^2 dr$, and each of these particles contributes $\mathbf{r}_i \cdot \nabla_i V_N(\mathbf{r}^N) = rv'(r)$ where $v' = dv/dr$. The result must be multiplied by $1/2$ so as to count each pair interaction only once. Thus we have

$$\left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \nabla_i V_N(\mathbf{r}^N) \right\rangle = 2\pi N\rho \int r^3 g(r)v'(r)dr. \quad (2.56)$$

By inserting this in Eq. (2.55) we obtain

$$\frac{\beta p}{\rho} = 1 - \frac{2\pi\beta\rho}{3} \int_0^\infty v'(r)g(r)r^3 dr. \quad (2.57)$$

For hard spheres, the pair interaction is infinite if $r < \sigma$ and zero otherwise. In this case the pressure equation cannot be used directly because its derivative cannot be computed. The way around is by introducing the so-called indirect correlation function $y(r) = \exp(\beta v(r))g(r)$ so that $v'(r)g(r) = -\beta^{-1}y(r)d \exp(-\beta v(r))/dr$. For hard-sphere $v(r)$, $d \exp(-\beta v(r))/dr = \delta(r - \sigma)$ and we have

$$\frac{\beta p}{\rho} = 1 + \frac{2\pi\rho}{3} \int_0^\infty y(r)\delta(r - \sigma)r^3 dr \quad (2.58)$$

$$= 1 + \frac{2\pi\rho}{3}\sigma^3 \lim_{r \rightarrow \sigma^+} y(r) \quad (2.59)$$

$$= 1 + \frac{2\pi\rho}{3}\sigma^3 g(\sigma). \quad (2.60)$$

The pressure of the hard-sphere liquid is determined by the value of the radial distribution function at contact.

In the grand-canonical potential, N is not fixed so that the normalization of the n -particle densities [Eq. (2.32)] reads

$$\int \rho_N^{(n)}(\mathbf{r}^n) d\mathbf{r}^n = \left\langle \frac{N!}{(N-n)!} \right\rangle \quad (2.61)$$

which gives

$$\int \rho_N^{(1)}(\mathbf{r}) d\mathbf{r} = \langle N \rangle \quad (2.62)$$

and

$$\int \rho_N^{(2)}(\mathbf{r}^{(2)}) d\mathbf{r}_1 d\mathbf{r}_2 = \langle N(N-1) \rangle = \langle N^2 \rangle - \langle N \rangle. \quad (2.63)$$

In the grand-canonical formalism, compressibility is defined by

$$\rho k_B T \chi_T = \frac{\langle N^2 \rangle - \langle N \rangle^2}{\langle N \rangle}. \quad (2.64)$$

The right-hand side can be expressed in terms of Eqs. (2.62) and (2.63) to give

$$\rho k_B T \chi_T = 1 + \frac{1}{\langle N \rangle} \int \left[\rho_N^{(2)}(\mathbf{r}_1, \mathbf{r}_2) - \rho_N^{(1)}(\mathbf{r}_1) \rho_N^{(1)}(\mathbf{r}_2) \right] d\mathbf{r}_1 d\mathbf{r}_2. \quad (2.65)$$

Now in a homogeneous system $\rho^{(1)} = \rho$ and $\rho^{(2)} = g^{(2)}\rho^2$ [cf. Eq. (2.40)] where $g^{(2)}$ depends only on r . Integrating over \mathbf{r}_1 gives V and $V/\langle N \rangle$ cancels one of the ρ in the ρ^2 present in both terms. The final result is the compressibility equation of state

$$\rho k_B T \chi_T = 1 + \rho \int \left[g^{(2)}(\mathbf{r}) - 1 \right] d\mathbf{r}. \quad (2.66)$$

If $g(r)$ is exact, the energy, pressure, and compressibility equations of state lead to the same results. Thus the derivation of the equation of state from a given pair potential $v(r)$ has been reduced to the search of the radial distribution function $g(r)$.

2.4 Distribution function theories

We first define the pair correlation function by

$$h(r) = g(r) - 1. \quad (2.67)$$

At small separations where $g(r)$ virtually vanishes $h(r) = -1$ — this is the anticorrelation hole. At the nearest-neighbor peak $h(r) > 0$ whereas at large separations where $g(r) = 1$ and $h(r) = 0$. So the pair correlation function makes sense.

The Ornstein-Zernike equation is derived by considering the response of the liquid of particles interacting by $v(r)$ to an external field. We skip the derivation because the result is quite intuitive — the relation between the total pair correlation function $h(r)$ and the direct correlation function $c(r)$ reads

$$h(r) = c(r) + \rho \int c(|\mathbf{r} - \mathbf{r}'|)h(r')d\mathbf{r}'. \quad (2.68)$$

This equation states that the total correlation function includes both direct correlations and correlations mediated by all other particles located at r' (Fig. 2.3). The probability of finding the other particles at r' is proportional to ρ , which is why it appears as a prefactor.

The effect of indirect correlations is best seen if the integral equation is solved recursively:

$$h(r) = c(r) + \rho \int c(|\mathbf{r} - \mathbf{r}'|)c(r')d\mathbf{r}' + \rho^2 \int c(|\mathbf{r} - \mathbf{r}'|)c(|\mathbf{r}' - \mathbf{r}''|)c(r'')d\mathbf{r}'d\mathbf{r}'' + \dots \quad (2.69)$$

The Ornstein-Zernike equation provides a different perspective of the pair correlations compared to the Yvon-Born equation. It can be considered a defining equation for $c(r)$ which is assumed to be related to the pair potential in some simple way. The pair potential $v(r)$ does not enter the equation explicitly. The problem is that we have a single equation for two unknown functions [$h(r)$ and $c(r)$]. An additional closure relation is needed to solve the system. Closure relations are obtained by approximating the exact $g(r)$ or $c(r)$ using graph-theoretical or density-functional methods. This leads to thermodynamic inconsistency: The equations of state derived from the energy, pressure, and compressibility equations are not the same.

A complete theoretical derivation of the various approximations is beyond our scope. Here we merely list the main features and the results.

Mean-spherical approximation

This is the simplest approximation of all. We assume that the density is small so that the second term in the Ornstein-Zernike equation is small compared to the first one so that:

$$h(r) \approx c(r). \quad (2.70)$$

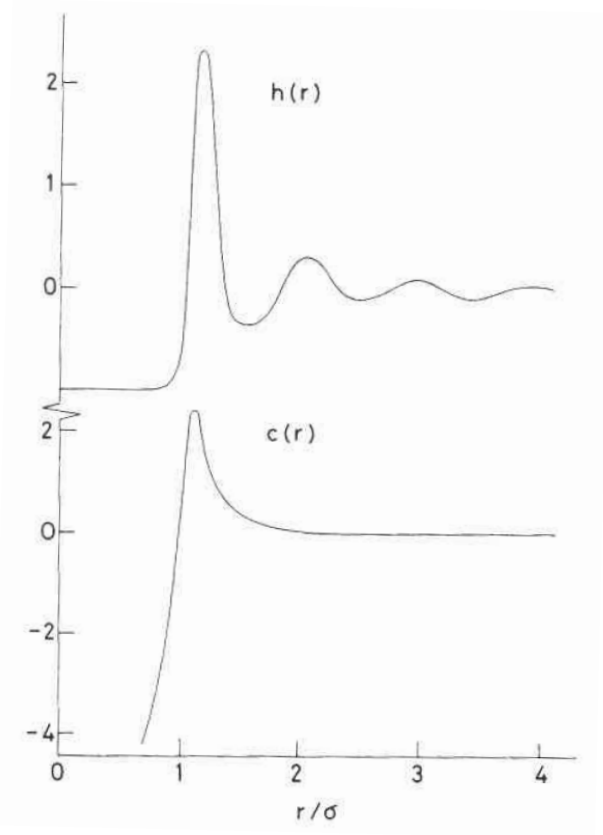


Figure 2.3: Total correlation function $h(r)$ and direct correlation $c(r)$ [3].

Now $h(r) = g(r) - 1$ which can be approximated by $\exp(-\beta v(r)) - 1 \approx -\beta v(r)$ if we assume that the pair interaction is weak. Thus

$$c(\mathbf{r}) = -\beta v(\mathbf{r}). \quad (2.71)$$

Clearly this only makes sense if $|\beta v(\mathbf{r})| \ll 1$ otherwise $g(r)$ may be negative which is non-physical.

Percus-Yevick approximation

We rewrite the Ornstein-Zernike equation:

$$c(r) = h(r) - \rho \int c(|\mathbf{r} - \mathbf{r}'|) h(r') d\mathbf{r}' = h(r) - h_{indirect}(r) = g(r) - g_{indirect}(r) \quad (2.72)$$

where we have introduced the indirect total correlation function $h_{indirect}(r)$. The indirect radial distribution function $g_{indirect}(r)$ can be expressed in terms of the effective potential $\Phi(r)$. The effective potential defines the radial distribution function $g(r)$ and is introduced based on the notion that at small densities, $g(r) \approx \exp(-\beta v(r))$ [Eq. (2.41)]. Thus we assume that as the density is increased, the functional form remains the same but $v(r)$ is replaced by an effective potential. This implies that the indirect correlation function is given by the difference between the effective and the pair potentials:

$$g_{indirect}(r) = \exp(-\beta[\Phi(r) - v(r)]) = g(r) \exp(\beta v(r)). \quad (2.73)$$

After inserting this result into Eq. (2.72) we have

$$c(r) = g(r) [1 - \exp(\beta v(r))]. \quad (2.74)$$

To lowest order, this is the same as Eq. (2.71) as $g(r) \approx \exp(-\beta v(r)) \approx 1 - \beta v(r)$.

Hypernetted chain approximation

If the pair potential $v(r)$ is weak, the exponential in Eq. (2.73) can be expanded $g_{indirect}(r) = \exp(-\beta[\Phi(r) - v(r)]) \approx 1 - \beta[\Phi(r) - v(r)] = 1 - \ln g(r) + \beta v(r)$ and

$$c(r) = g(r) - g_{indirect}(r) = g(r) - 1 + \ln g(r) - \beta v(r). \quad (2.75)$$

To lowest order, this too is the same as Eq. (2.71) since $g(r) \approx 1 - \beta v(r)$.

2.5 Virial expansion

The main assumption of the virial expansion is that the equation of state can be written as a power series of the density. While this is reasonable at small densities, the expansion becomes more and more impractical as the density is increased.

The virial expansion reads

$$\frac{\beta p}{\rho} = 1 + \sum_{i=2}^{\infty} B_i(T) \rho^{i-1}, \quad (2.76)$$

where B_i are the virial coefficients. The derivation of the coefficients is complicated but B_2 can be obtained from the pressure equation of state Eq. (2.57). By approximating $g(r)$ by $\exp(-\beta v(r))$ we have

$$\frac{\beta p}{\rho} = 1 - \frac{2\pi\beta\rho}{3} \int_0^\infty v'(r) \exp(-\beta v(r)) r^3 dr. \quad (2.77)$$

The second term is now integrated by parts [$r^3 = u$, $-\beta v' \exp(-\beta v(r)) dr = dv$] and we obtain

$$\frac{\beta p}{\rho} = 1 - 2\pi\rho \int_0^\infty [\exp(-\beta v(r)) - 1] r^2 dr \quad (2.78)$$

so that

$$B_2(T) = -\frac{1}{2} \int f(\mathbf{r}) d\mathbf{r} \quad (2.79)$$

where $f(\mathbf{r}) = \exp(-\beta v(r)) - 1$ is the Mayer function.

The other virial coefficients are defined by irreducible Mayer cluster integrals β_i [3]. B_3 represents triplets:

$$B_3 = -\frac{1}{3} \int f(r)f(r')f(|\mathbf{r} - \mathbf{r}'|) d\mathbf{r}d\mathbf{r}'. \quad (2.80)$$

2.6 Hard-sphere equation of state*

This is a very important reference system because it accounts for the impenetrability of the molecules. The attractive interaction between the molecules can be incorporated perturbatively as discussed in the next section.

For hard-sphere potential

$$v(r) = \begin{cases} \infty & r < \sigma \\ 0 & r > \sigma \end{cases}. \quad (2.81)$$

The Percus-Yevick direct correlation function is [3]

$$c(r/\sigma) = \begin{cases} -\lambda_1 - 6\eta\lambda_2 r/\sigma - \eta\lambda_1(r/\sigma)^3/2 & r < \sigma \\ 0 & r > \sigma \end{cases}, \quad (2.82)$$

where

$$\eta = \frac{\pi\rho\sigma^3}{6} \quad (2.83)$$

is the packing fraction whereas $\lambda_1 = (1 + 2\eta)^2/(1 - \eta)^3$ and $\lambda_2 = -(1 + \eta/2)^2/(1 - \eta)^4$. After inserting this into the compressibility equation we find

$$\frac{\beta p}{\rho} = \frac{1 + \eta + \eta^2}{(1 - \eta)^3}. \quad (2.84)$$

On the other hand, inserting this into the pressure equation of state gives

$$\frac{\beta p}{\rho} = \frac{1 + 2\eta + 3\eta^2}{(1 - \eta)^3}. \quad (2.85)$$

These two equations are not the same.

Carnahan-Starling equation of state*

In hard spheres, the second virial coefficient B_2 is trivial

$$B_2 = -\frac{1}{2} \int f(r) \mathbf{d}\mathbf{r} = 2\pi \int_0^\sigma r^2 \mathbf{d}r = \frac{2\pi\sigma^3}{3}. \quad (2.86)$$

The third virial coefficient B_3 is less trivial (homework) and the result is:

$$B_3 = \frac{5\pi^2\sigma^6}{18}. \quad (2.87)$$

The fourth virial coefficient reads

$$B_4 = \left(-\frac{89}{280} + \frac{219\sqrt{2}}{2240\pi} + \frac{4131}{2240\pi} \arccos \frac{1}{\sqrt{3}} \right) B_2^2 = 0.28695 B_2^2. \quad (2.88)$$

$B_5, B_6,$ and B_7 were computed numerically.

Carnahan and Starling expressed the virial equation of state by replacing ρ in the series expansion by $\eta = \pi\rho\sigma^3/6$. The coefficients are rescaled too:

$$\mathcal{B}_i = \left(\frac{6}{\pi\sigma^3} \right)^i B_{i+1}. \quad (2.89)$$

This gives

$$\frac{\beta p}{\rho} = 1 + 4\eta + 10\eta^2 + 18.365\eta^3 + 28.24\eta^4 + 39.5\eta^5 + 56.5\eta^6 + \dots \quad (2.90)$$

Then they noticed that if the rescaled coefficients are rounded to the 4, 10, 18, 28, 40, and 54, they can be expressed by a simple formula

$$\mathcal{B}_i = i^2 + 3i, \quad (2.91)$$

where $i = 1, 2, 3, 4, 5,$ and 6 . The advantage of this closed expression for the coefficients is that if we assume that all \mathcal{B}_i are given by this formula the virial equation of state can be summed:

$$\frac{\beta p}{\rho} = 1 + \sum_{i=1}^{\infty} (i^2 + 3i)\eta^i. \quad (2.92)$$

The two terms are related to the second and the first derivative of the geometric series and the final result is

$$\frac{\beta p}{\rho} = \frac{1 + \eta + \eta^2 - \eta^3}{(1 - \eta)^3}. \quad (2.93)$$

This equation of state agrees well with the numerical results at packing fractions where the liquid is stable (Fig. 2.4).

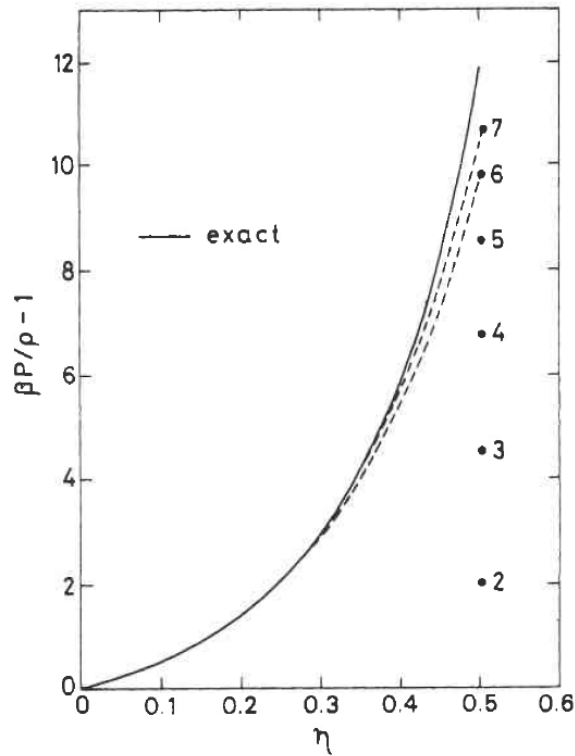


Figure 2.4: Comparison of the Carnahan-Starling hard-sphere equation of state (solid line) and the virial series including 2, 3, ... 7 terms [3].

2.7 Perturbation theories

In most cases the intermolecular potential consists of two fairly distinct parts. The overlap of electron clouds leads to short-range repulsion whereas the attraction is typically more long-ranged. The repulsive part of the potential is responsible for the structure of the liquid and the attraction provides the cohesive background potential.

This splitting suggests that it is useful to consider splitting the intermolecular potential into a reference part (e.g., the hard-core potential) and the remainder which is then treated as a perturbation. In the van der Waals model, the excluded volume associated with each molecule of diameter σ is

$$V_{excl} = \frac{1}{2} \frac{4\pi\sigma^3}{3} = \frac{2\pi\sigma^3}{3}. \quad (2.94)$$

The prefactor of $1/2$ appears because the excluded volume can only make sense in pairs of molecules, and the rest is the volume of a sphere of radius σ around each molecule. If we replace V in the ideal-gas equation of state by $V - V_{excl}$ it reads

$$\frac{\beta p}{\rho} = \frac{1}{1 - 4\eta}. \quad (2.95)$$

This formula provides the correct B_2 and simultaneously suggests that the virial expansion may not make a lot of sense. On the other hand, the pressure diverges at $\eta = 1/4$ which is way too low — in hard spheres, freezing begins at $\eta = 0.49$.

In the λ -expansion, the potential is written as

$$v_\lambda(r) = v_0(r) + \lambda w(r), \quad (2.96)$$

where $\lambda_0 = 0$ gives the reference potential $v_0(r)$ and $\lambda_1 = 1$ gives the desired potential. We assume that the total potential energy is given by a sum of pair potentials

$$V_N(\lambda) = \sum_{i < j} v_\lambda(r_{ij}). \quad (2.97)$$

The derivative of the free energy with respect to λ reads

$$\beta \frac{\partial F}{\partial \lambda} = \frac{1}{Z_N(\lambda)} \int \exp(-\beta V_N(\lambda)) \beta V'_N(\lambda) d\mathbf{r}^N = \beta \langle V'_N(\lambda) \rangle. \quad (2.98)$$

where $V'_N(\lambda) = \partial V_N(\lambda) / \partial \lambda$. After this derivative is evaluated, we can integrate it to calculate the free energy

$$\beta F(\lambda_1) = \beta F(\lambda_0) + \beta \int_{\lambda_0}^{\lambda_1} \langle V'_N(\lambda) \rangle d\lambda. \quad (2.99)$$

In the simplest case where the intermolecular potential is pairwise additive, the free energy per molecule reads

$$\frac{\beta F}{N} = \frac{\beta F_0}{N} + \frac{\beta}{2N} \int_0^1 d\lambda \int \rho_{\lambda}^{(2)}(\mathbf{r}_1, \mathbf{r}_2) w(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2. \quad (2.100)$$

The prefactor 1/2 in front of the integral is due to pair interaction. We have replaced λ_0 by 0 and λ_1 by 1; F_0 denotes $F(\lambda = 0)$.

Now the pair density depends on λ but if it is replaced by $\rho_{\lambda=0}^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$ we have

$$\frac{\beta F}{N} = \frac{\beta F_0}{N} + \frac{\beta}{2N} \int_0^1 d\lambda \int \rho_{\lambda=0}^{(2)}(\mathbf{r}_1, \mathbf{r}_2) w(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2. \quad (2.101)$$

Given that $\rho_{\lambda=0}^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \rho^2 g_0(r)$ and that the integration over $d\mathbf{r}_1$ gives V as the system is homogeneous we have

$$\frac{\beta F}{N} = \frac{\beta F_0}{N} + \frac{\beta}{2N} \rho^2 V \int_0^1 d\lambda \int g(r) w(r) d\mathbf{r} \quad (2.102)$$

$$= \frac{\beta F_0}{N} + \frac{\beta \rho}{2} \int g(r) w(r) d\mathbf{r}. \quad (2.103)$$

The integral over λ is trivial, and \mathbf{r} is the position of particle 2 relative to particle 1.

If the range of the pair potential is much longer than the distance at which $g(r)$ approaches the asymptotic value of 1 (which depends on density), the remaining integral can be approximated by

$$\frac{\beta \rho}{2} \int g(r) w(r) d\mathbf{r} \approx \frac{\beta \rho}{2} \int w(r) d\mathbf{r} = -\beta \rho a, \quad (2.104)$$

where a is a constant and $a > 0$ for attractive pair interaction $w(r)$.

Now we compute the equation of state. From

$$p = - \left(\frac{\partial F}{\partial V} \right)_T = \frac{\rho^2}{N} \left(\frac{\partial F}{\partial \rho} \right)_T \quad (2.105)$$

and

$$\frac{\beta p}{\rho} = \frac{\beta \rho}{N} \left(\frac{\partial F}{\partial \rho} \right)_T = \rho \left(\frac{\partial(\beta F/N)}{\partial \rho} \right)_T \quad (2.106)$$

we have

$$\frac{\beta p}{\rho} = \frac{\beta p_0}{\rho} - \beta \rho a, \quad (2.107)$$

where

$$\frac{\beta p_0}{\rho} = \rho \left(\frac{\partial(\beta F_0/N)}{\partial \rho} \right)_T \quad (2.108)$$

is the equation of state of the reference state.

If the reference state is the hard-core liquid modeled by Eq. (2.95) and if Eq. (2.95) is inserted in Eq. (2.107) we have

$$\frac{\beta p}{\rho} = \frac{1}{1 - 4\eta} - \beta \rho a. \quad (2.109)$$

This can be rearranged to give

$$\frac{\beta (p + a\rho^2)}{\rho} = \frac{1}{1 - 4\eta} \quad (2.110)$$

$$(p + a\rho^2) (\rho^{-1} - 4\eta/\rho) = \beta^{-1}. \quad (2.111)$$

As $\eta/\rho = \pi\sigma^3/6$, this result is of the well-known van der Waals equation of state

$$\left(p + \frac{\tilde{a}}{V_M^2} \right) (V_M - b) = RT, \quad (2.112)$$

where $\tilde{a} = N_A^2 a$ and $b = 2\pi N_A \sigma^3/3$ are constants and $V_M = MV/m$ is the kilomolar volume.

Chapter 3

Liquid crystals

The traditional division of states of matter into crystals of long-range 3D positional order (and perfect orientational order in case of molecular crystals) and simple liquids with no long-range positional order barely scratches the surface of the phenomenology of condensed matter. There exist a range of mesophases characterized by various types of intermediate positional and/or orientational order, and these mesophases are usually referred to as liquid crystals. The term liquid crystals describes the liquid-like response to shear (most liquid crystals flow like simple liquids) and their crystal-like optical properties, especially birefringence.

Liquid crystals are materials based on anisotropic molecules. The molecular anisotropy leads to orientational order described by a headless vector called the **director** \mathbf{n} , which specifies the average molecular orientation. The most common types of liquid crystals are

- **Nematics:** Characterized by orientational order only and based on elongated rod-like molecules, nematics are the most disordered liquid crystals. Nematic order based on chiral molecules (which differ from their mirror image much like the left and the right screw) is chiral too; these phases are traditionally called cholesteric. In chiral nematic phase, the director twists around a perpendicular axis and the pitch is typically comparable to the wavelength of light (Fig. 3.1).
- **Smectics:** These phases are also found in rod-like molecules and have a partial positional order in addition to the orientational order. The molecules are arranged in layers which are 2D liquid. In smectic A (Sm A) phase, the director is perpendicular to the layers whereas in smectic C (Sm C) phase it is tilted with respect to layer normal. There exist many subtypes of smectic phases which include chiral phases.

- **Columnar liquid crystals:** These are typically found in plate-like particles which are arranged in columns much like coins. There is no long-range positional order within the columns but the particles are orientationally ordered. In addition, there exists 2D positional order of the columns.

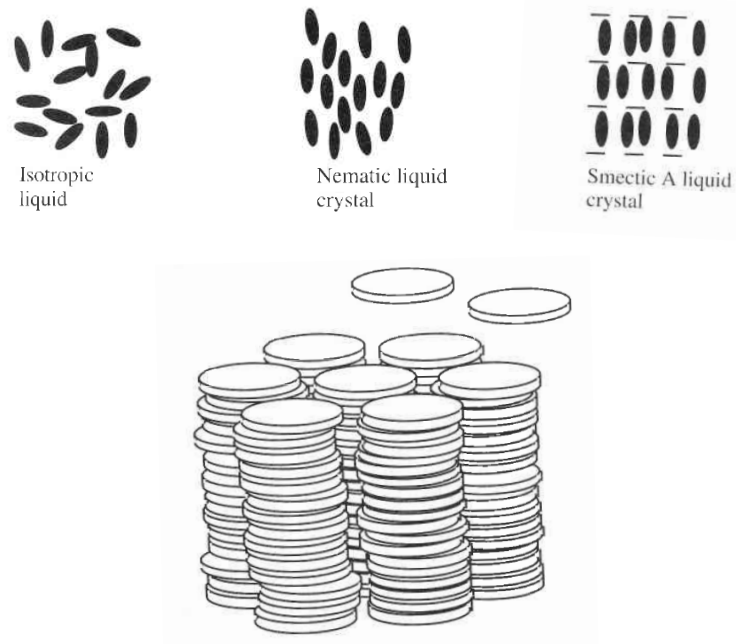


Figure 3.1: Schematic of the simplest liquid-crystalline phases: Isotropic, nematic, smectic, and columnar phase [1].

Most liquid crystal phases are found in organic molecules of suitable shape. In these materials usually referred to as thermotropic liquid crystals phase transitions are induced by temperature. Liquid-crystal order also exists in anisotropic colloidal particles. Especially well-studied are liquid-crystalline phases in tobacco mosaic and fd virus (Dogic and Fraden) and in various kinds of clay particles (Lekkerkerker).

From the conceptual point of view, it is important to understand how does the anisotropic shape of either molecules or colloids result in liquid-crystalline order. The basic insight into the microscopic structure of liquid crystals is provided by the Onsager theory of nematic-isotropic transition.

3.1 Onsager theory

In the Onsager theory, one considers a system of hard rods of length L and diameter D [1, 15]. The assumptions made are: i) the interaction between the rods is steric; ii) the packing fraction $\eta = (N/V)\pi LD^2/4 = \pi LD^2\rho/4$ is small, and iii) the aspect ratio of the rods is large, $L/D \gg 1$.

The free energy of a suspension of rods is constructed based on what we know about ideal gases and gases of finite-size spheres. In ideal gases the configuration integral is $Z_N(V, T) = V^N$ whereas in dilute finite-size spheres $Z_N(V, T) \approx V_{free}^N$ where $V_{free} = V - NV_{excl}$ is the free volume available to the particles; here $V_{excl} = 2\pi\sigma^3/3$ is the excluded volume [Eq. (2.94)]. To lowest order, the excess free energy of a gas of finite-size spheres reads

$$F^{ex} = -k_B T \ln \left(\frac{Z_N(V, T)}{V^N} \right) = -k_B T \ln \left(1 - \frac{NV_{excl}}{V} \right)^N \approx Nk_B T \frac{NV_{excl}}{V}. \quad (3.1)$$

In hard rods, the combined *positional* ideal and excess free energy are

$$F = F^{id} + F^{ex} = -Nk_B T \ln V + Nk_B T \frac{NV_{excl}}{V}. \quad (3.2)$$

The appearance of orientational order gives rise to the dependence of the excluded volume V_{excl} on the relative orientation of the particles. The excluded volume of parallel rods is a cylinder of radius D and length L totaling to $\pi D^2 L$ (Fig. 3.2). For perpendicular rods, the excluded volume is a cylinder of ellipsoidal base (Fig. 3.2). The excluded volume can be straightforwardly calculated in the limit of large L/D . In this limit, the dominant contribution comes from rods forming a Y or T shape (so that the tip of one rod bumps into the other rod) where the center-to-center distance is of the order $L/2$ and the excluded volume of the crossed X shaped configurations where the distance between the rod axes is D can be neglected. In this approximation, the excluded volume consists of two rhombic prisms of height $2D$ and base area $L^2 |\sin \gamma|/2$ where γ is the angle between the long axes (Fig. 3.2). This volume is to be divided by 2 so as to associate half of it to each particle. The final excluded volume per particle is

$$V_{excl} = DL^2 |\sin \gamma|. \quad (3.3)$$

If the orientation of the rods $f(\theta)$ is isotropic, the average $|\sin \gamma| = \pi/2$. The average value of $|\sin \gamma|$ is a functional of the orientational distribution $f(\theta)$

$$p[f(\theta)] = \langle |\sin \gamma| \rangle = \int f(\theta) f(\theta') \sin \gamma \, d\Omega d\Omega', \quad (3.4)$$

where Ω and Ω' describe the orientations of the two particles in space.

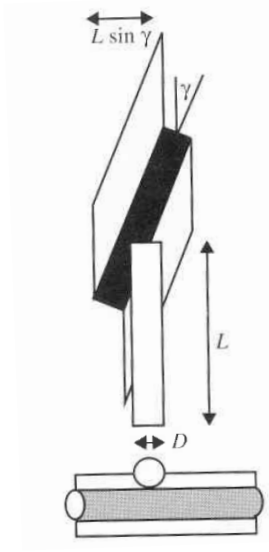


Figure 3.2: Onsager theory: rod-rod excluded volume [1].

Secondly, the orientational order decreases the entropy. Using the Gibbs formula, we can express the entropy loss due to the anisotropic distribution of rod orientations, and the corresponding excess free energy difference is

$$\Delta F^{ori} = Nk_B T \int f(\theta) \ln(4\pi f(\theta)) d\Omega. \quad (3.5)$$

Thus the total free energy reads

$$F = F^{id} + F^{ex} \quad (3.6)$$

$$= Nk_B T \left\{ -\ln V + DL^2 \rho p[f(\theta)] + \int f(\theta) \ln(4\pi f(\theta)) d\Omega \right\} \quad (3.7)$$

$$= Nk_B T \left\{ \ln \rho + DL^2 \rho p[f(\theta)] + \int f(\theta) \ln(4\pi f(\theta)) d\Omega \right\} + const. \quad (3.8)$$

In terms of the packing fraction $\eta = \rho\pi LD^2/4$, $DL^2\rho$ reads $L\eta/D$; η is also used to rewrite $\ln \rho$ as $\ln(L\eta/D) - \ln(D/L) = \ln(L\eta/D) + const.$ This gives the final form of the Onsager free energy

$$F = Nk_B T \left\{ \ln \frac{L\eta}{D} + \frac{4}{\pi} \frac{L\eta}{D} p[f(\theta)] + \int f(\theta) \ln(4\pi f(\theta)) d\Omega \right\} + const. \quad (3.9)$$

The only parameter of this theory is $L\eta/D$.

We will solve the problem variationally using an ansatz for the orientational distribution $f(\theta)$. A reasonable normalized ansatz is

$$f(\theta) = \frac{\alpha \cosh(\alpha \cos \theta)}{4\pi \sinh \alpha}. \quad (3.10)$$

The nematic phase corresponds to large $\alpha \sim 20$ where the distribution is strongly peaked at $\theta = 0$ and $\theta = \pi$. For this particular choice of $f(\theta)$ we can compute the degree of orientational order (often called the nematic order parameter) defined by

$$S = \langle P_2(\cos \theta) \rangle \quad (3.11)$$

$$= \frac{1}{2} \int f(\theta)(3 \cos^2 \theta - 1) d\Omega \quad (3.12)$$

$$= \frac{1}{2} \int f(\theta)(3 \cos^2 \theta - 1) 2\pi \sin \theta d\theta, \quad (3.13)$$

where $P_2(x) = (3x^2 - 1)/2$ is the Legendre polynomial of order 2. The result is

$$S(\alpha) = 1 - 3 \frac{\coth \alpha}{\alpha} + \frac{3}{\alpha^2}. \quad (3.14)$$

S is a monotonically increasing function of α , starting from 0 at $\alpha = 0$ and saturating at 1 for $\alpha \rightarrow \infty$. For $\alpha \ll 1$, $S \approx \alpha^2/15$ whereas for $\alpha \gg 1$, $S \approx 1 - 3/\alpha$.

Using the ansatz Eq. (3.10) we can evaluate the average $|\sin \gamma|$ and plot the Onsager free energy [Eq. (3.9)] vs. S for various $L\eta/D$. As α and thus S are increased, the orientational free energy $Nk_B T \int f(\theta) \ln(4\pi f(\theta)) d\Omega$ increases — the larger the order parameter, the more peaked the orientational distribution and the smaller the entropy. On the other hand, $p[f(\theta)]$ tends to 0 as S goes to 1: If all molecules are parallel to each other, the excluded volume is very small (zero in the limit $L/D \rightarrow \infty$). The weight of the excluded-volume term is $4L/(\pi D)$ so that at vanishingly small L/D the free energy is minimized by the isotropic state with $S = 0$ whereas at large enough L/D there appears a minimum at a finite S .

The Onsager free energy is shown in Fig. 3.3. We see that for $L\eta/D < 3.34$, free energy is minimal at $S = 0$ which corresponds to the isotropic phase whereas for $L\eta/D > 4.49$ the minimum is at a finite $S > 0$. For $L\eta/D$ between these two values, isotropic phase coexists with a nematic phase with $S = 0.84$.

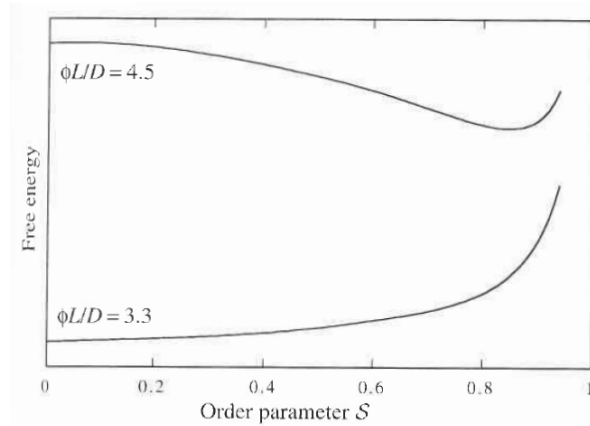


Figure 3.3: Onsager theory: Free energy as a function of S for $L\eta/D = 3.3$ (isotropic phase) and $L\eta/D = 4.4$ (nematic phase) [1].

Based on these results, we can plot the phase diagram in the $(\eta, L/D)$ -plane (Fig. 3.4). In the limit of very big packing fractions $\eta \rightarrow 1$, nematic phase is stable for $L/D > 4.5$.

In rods, the maximal achievable packing fraction is 0.907 and the packing fraction of nematics is somewhat smaller than this but still of the same order of magnitude. We can conclude that an aspect ratio of about 5 is sufficient to stabilize the nematic phase in dense systems such as most molecular liquid crystals. On the other hand, very long particles such as fd or tobacco mosaic viruses form colloidal nematic phases at small packing fractions.

A more detailed insight into the role of particle shape can be obtained using numerical simulations. In an analysis of hard spherocylinders [6], nematic phase is found at $L/D > 4.6$. As far as the nematic-isotropic transition is concerned, the phase diagram is consistent with the predictions of the Onsager theory (Fig. 3.5; note that the x and y axes in this figure are swapped compared to Fig. 3.4). Also present in the phase diagram are the smectic and the crystalline phases.

An important lesson learned from the Onsager theory and from the numerical studies of systems of anisotropic hard-core particles is that the anisotropy of molecular shape alone is sufficient for the stabilization of the orientationally ordered phases. This is also witnessed by the form of the Onsager free energy [Eq. (3.9)] which is athermal: Temperature only enters as a prefactor, emphasizing that the origin of the orientational order is entropic. This does not mean that the interactions between the liquid-crystalline molecules

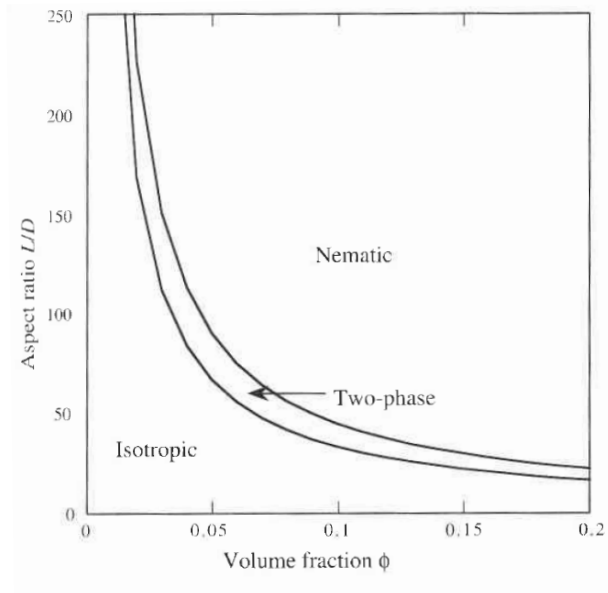


Figure 3.4: Phase diagram of the Onsager theory [1].

or colloids is entirely hard-core; but the steric repulsion between them is the reference interaction.

3.2 Nematic elasticity*

The advantage of statistical-mechanical theories such as the Onsager theory is that they provide an insight into the workings of the liquid-crystalline order. In most liquid-crystalline samples and applications, the order is spatially modulated rather than uniform like in a well-ordered bulk sample. The modulations may be imposed by external forces such as aligning surfaces. To describe the spatially distorted order, one needs to introduce an elastic theory associated to a suitable order parameter.

The most obvious order parameter is the nematic director \mathbf{n} and the most widely used theory of elasticity is due to Frank [15]. In the Frank theory, we consider spatial distortions of the director $\mathbf{n}(\mathbf{r})$. These distortions are very gradual so that $L\nabla\mathbf{n} \ll 1$ where L is the length of the molecules. The assumptions made are the following:

- i) the distortion free energy per unit volume f_d can be expanded in powers of $\nabla\mathbf{n}$ so that in a uniform nematic f_d vanishes;

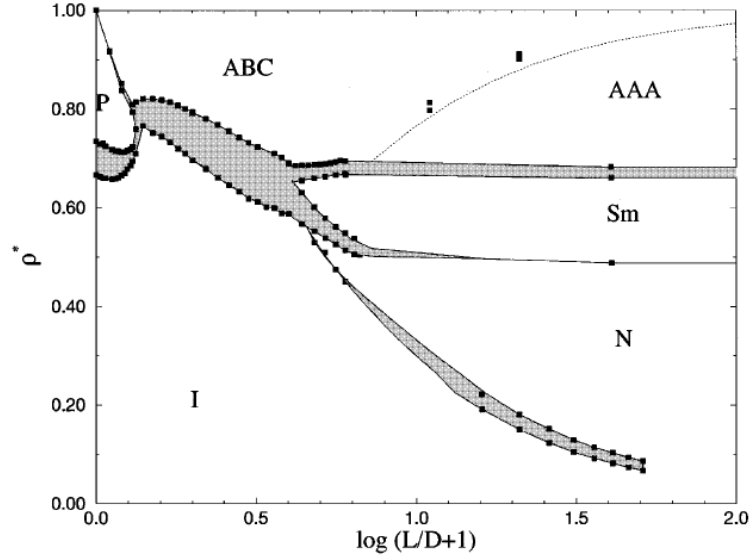


Figure 3.5: Phase diagram of hard spherocylinders [6]. $L + D$ and D are the spherocylinder length and diameter, respectively, so that $L/D + 1$ is the aspect ratio.

- ii) f_d is even in \mathbf{n} because \mathbf{n} and $-\mathbf{n}$ are indistinguishable;
- iii) f_d does not contain terms linear in $\nabla\mathbf{n}$ — the only two terms of this type allowed are $\nabla \cdot \mathbf{n}$ (which changes sign as \mathbf{n} is replaced by $-\mathbf{n}$ and is thus forbidden) and $\mathbf{n} \cdot \nabla \times \mathbf{n}$ (which changes sign for $\mathbf{r} \rightarrow -\mathbf{r}$ and is thus forbidden in materials with center of symmetry);
- iv) there are no terms of the form $\nabla\mathbf{u}$ where \mathbf{u} is an arbitrary vector (because these terms can be transformed into surface integrals).

The derivation of all terms allowed by symmetry is somewhat involved; the result is

$$f_d = \frac{K_{11}}{2} (\nabla \cdot \mathbf{n})^2 + \frac{K_{22}}{2} (\mathbf{n} \cdot \nabla \times \mathbf{n})^2 + \frac{K_{33}}{2} (\mathbf{n} \times \nabla \times \mathbf{n})^2. \quad (3.15)$$

The three terms represent the splay, twist, and bend deformation, respectively (Fig. 3.6), whereas K_{11} , K_{22} , and K_{33} are the corresponding elastic constants. The magnitude of all $K_{ii} \sim k_B T / L \approx 10^{-12}$ N ($k_B T = 4 \times 10^{-21}$ J, $L \approx 10^{-10}$ m is the length of molecules). As the elastic constants are typically not very different from one another, they are sometimes

assumed to be equal which leads to the so-called one-constant approximation

$$f_d = \frac{K}{2} [(\nabla \cdot \mathbf{n})^2 + (\nabla \times \mathbf{n})^2]. \quad (3.16)$$

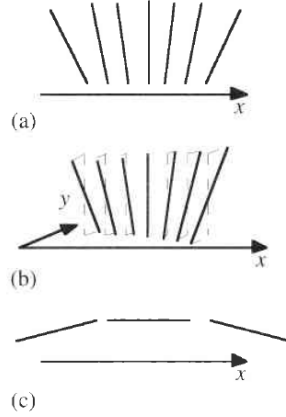


Figure 3.6: Splay, twist, and bend deformation [1].

To better understand the spatial nature of the three deformation modes, consider three examples of director fields slightly deviating from the uniform field along the x axis (Fig. 3.6). A simple example of the **splay** deformation is

$$\mathbf{n}(x) = (ax, 0, 1), \quad (3.17)$$

where $a \ll 1$ so that $|\mathbf{n}| = 1$ to lowest order. Obviously $\nabla \cdot \mathbf{n} = a$. To visualize the **twist** deformation, consider the director field of type

$$\mathbf{n}(y) = (ay, 0, 1). \quad (3.18)$$

For this director field, $\nabla \times \mathbf{n} = (0, 0, -a)$ and the dot product $\mathbf{n} \cdot \nabla \times \mathbf{n} = -a$ directly reflects the deviation of the director from the uniform field. The **bend** deformation can be illustrated by

$$\mathbf{n}(z) = (az, 0, 1), \quad (3.19)$$

where again $a \ll 1$. This gives $\nabla \times \mathbf{n} = (0, a, 0)$ and the cross product $\mathbf{n} \times \nabla \times \mathbf{n} = (-a, 0, a^2z) \approx (-a, 0, 0)$.

Chiral nematics

Chiral nematics are characterized by a finite spontaneous twist due to the chirality of the molecules. The ground state will ideally have a given preferred twist q_0 everywhere so that the Frank free energy reads

$$f_d = \frac{K_{11}}{2} (\nabla \cdot \mathbf{n})^2 + \frac{K_{22}}{2} (\mathbf{n} \cdot \nabla \times \mathbf{n} - q_0)^2 + \frac{K_{33}}{2} (\mathbf{n} \times \nabla \times \mathbf{n})^2, \quad (3.20)$$

where $q_0 = 2\pi/\lambda$ and λ is the pitch.

The traditional notion of a chiral nematic is summarized by the helical director field $\mathbf{n}(z) = (\cos q_0 z, \sin q_0 z, 0)$. But this structure does not really produce the ideal twisted structure because the director only twists in one direction perpendicular to the director (z) but not in the other perpendicular direction (y). The ideal structure is found in the core of the double-twist cylinder (Fig. 3.7) but on going outwards from the center of the cylinder the double-twist structure is gradually violated. — The double-twist cylinders are the essential building block of the **blue phases**.

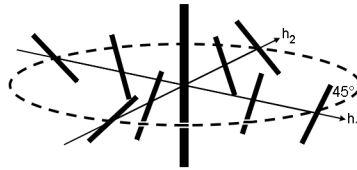


Figure 3.7: Double-twist structure. Source: Wikipedia, http://en.wikipedia.org/wiki/Blue_phase_mode_LCD.

Surface anchoring

Most liquid-crystal devices are based on thin cells where the nematic is sandwiched between two glass plates treated to induce a specific preferred orientation of the director often referred to as the *easy axis*. The surfaces can be treated mechanically, typically by rubbing, or chemically, and the effect is called *anchoring*. In homeotropic anchoring, the easy axis is normal to the plates; in planar anchoring it lies in the plane but there exists no specific in-plane orientation; and in homogeneous anchoring the in-plane orientation is specified.

The anchoring may be either strong where the director at the surface is prescribed or weak where the director may deviate from the easy axis

but the deviation is associated to an increase of the anchoring energy. The anchoring energy is usually modeled by the Rapini-Papoular energy

$$F_s = -\frac{W}{2} \int (\mathbf{n} \cdot \mathbf{k})^2 dA, \quad (3.21)$$

where W is the anchoring strength and \mathbf{k} is the easy axis. This energy is minimal if $\mathbf{n} = \mathbf{k}$.

The Frank elastic constant and the anchoring strength can be combined into the so-called *extrapolation length*

$$\Lambda = \frac{K}{W}. \quad (3.22)$$

If anchoring is strong, $\Lambda \rightarrow 0$.

The relevance of the extrapolation length can be appreciated by the following example. Consider a twisted nematic cell with homogeneous anchoring at the two walls so that the easy axis points along the x axis at $z = 0$ and makes an angle of θ_0 with the x axis at $z = h$. Assume that the director field is twisted

$$\mathbf{n}(z) = (\cos \theta(z), \sin \theta(z), 0) \quad (3.23)$$

and also assume that the anchoring at $z = 0$ is strong. The total elastic energy per unit area reads

$$\frac{F_d + F_s}{A} = \frac{K_{22}}{2} \int \left(\frac{d\theta}{dz} \right)^2 dz + \frac{W}{2} (\Theta - \theta_0)^2 + \text{const.}, \quad (3.24)$$

where $\Theta = \theta(z = h)$ and the Rapini-Papoular anchoring energy has been approximated by the lowest-order, parabolic term: $(\mathbf{n} \cdot \mathbf{k})^2 = \cos^2(\Theta - \theta_0)$ in Eq. (3.21) has been expanded in Taylor series $1 - (\Theta - \theta_0)^2$ and the constant 1 has been omitted. The Euler-Lagrange equation is $d^2\theta/dz^2 = 0$ so that θ is a linear function of z :

$$\theta(z) = \Theta \frac{z}{h}. \quad (3.25)$$

The boundary condition is $\partial f_s / \partial \theta|_{z=h} = \partial f_d / \partial \theta'|_{z=h}$ where f_s is area density of the surface energy and $\theta' = d\theta/dz$. If spelled out, $W(\Theta - \theta_0) = -K_{22}\Theta/h$ or

$$\Theta - \theta_0 = -\frac{\Lambda}{h}\Theta. \quad (3.26)$$

The solution to this equation is

$$\Theta = \frac{1}{1 + \Lambda/h} \theta_0. \quad (3.27)$$

If the extrapolation length is small compared to cell thickness, $\Lambda \ll h$, $\Theta = \theta_0$ and the actual orientation of the director at the $z = h$ wall coincides with the easy axis $\Theta = \theta_0$. This is the strong anchoring regime. On the other hand, if Λ is comparable to h , then Θ is smaller than θ_0 .

External field

The electric and magnetic properties of nematogenic molecules are anisotropic. As a result, an external electric and magnetic field orients the director, usually such that the director is parallel to the field. In case of magnetic field, the interaction energy per unit volume can be written as

$$f_{mag} = -\frac{\chi_a}{2\mu_0} (\mathbf{n} \cdot \mathbf{B})^2, \quad (3.28)$$

where $\chi_a = \chi_{\parallel} - \chi_{\perp}$ is the anisotropy of the magnetic susceptibility (χ_{\parallel} and χ_{\perp} are the susceptibilities along and perpendicular to \mathbf{n} , respectively), and \mathbf{B} is the magnetic field. If $\chi_a > 0$, director tends to be parallel to the director and the energy is lowest for $\mathbf{n} \parallel \mathbf{B}$. — The effect of the electric field is qualitatively very similar.

Frederiks transition

The basic element of liquid-crystal displays is a thin cell filled with a nematic. The walls of the cell are treated to induce a specific anchoring and covered with transparent conductive electrodes. The anchoring easy axis is chosen such that it does not coincide with the direction of the electric field. As long as the field is off, the director structure is dictated by the anchoring only. After the field is turned on the director is reoriented and this changes the optical properties of the cell, typically such that the off state is transparent and the on state is dark.

The transition between the two states is continuous and takes place at a finite field. In the strong-anchoring cell, the critical field is inversely proportional to the cell thickness squared

$$B_c = \sqrt{\frac{\pi^2 \mu_0 K_{ii}}{\chi_a h^2}}. \quad (3.29)$$

The elastic constant in question depends on the type of deformation.

3.3 Defects*

The texture of liquid crystals is often nonuniform, showing regions where the director is obviously nonuniform (Fig. 3.8) [15]. On approaching the centers of these regions, the gradient of the director obviously diverges — this is witnessed by the rapid variation of the appearance and optical properties of the material. The points where the director field seems to be undefined are called *defects* or more correctly *disclinations*.

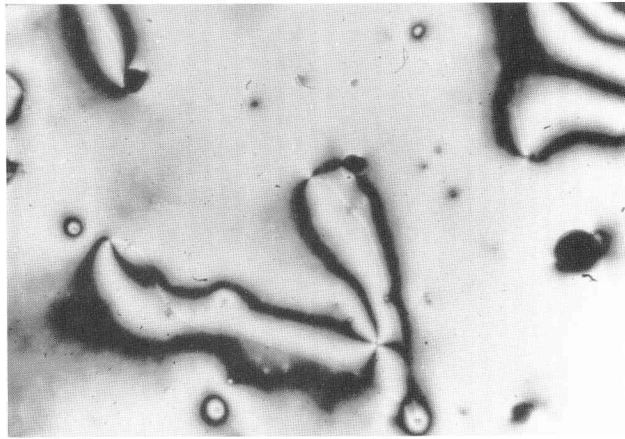


Figure 3.8: Defects in nematics as seen through crossed polarizers; this is the so-called Schlieren texture [15].

Imagine that the nematic director field is two-dimensional; this corresponds to *line defects*. Due to the equivalence of \mathbf{n} and $-\mathbf{n}$, the net change of the orientation of the director along a closed loop around the defect can be a half-integer multiple of 2π :

$$\oint \frac{d\theta}{ds} ds = 2\pi k, \quad (3.30)$$

where $k = \pm 1/2, \pm 1, \pm 3/2 \dots$ ($k = 0$ is also possible but this is a defect-free configuration.) The number k is the *defect strength*. A few typical defects are shown in Fig. 3.9. The $k = -1$ disclination is hyperbolic and does not rely on the $\mathbf{n} \leftrightarrow -\mathbf{n}$ equivalence; it exists in a proper vectorial field too. The $k = -1/2$ and $k = 1/2$ are typical for the nematic phase. The radial $k = 1$ and the annular $k = 1$ defect are of the same strength. In the former, the angle between the director and (say) x axis is given by $\theta = \phi$ where ϕ is

the azimuthal angle whereas in the latter $\theta = \phi + \pi/2$. Both director fields can be described by

$$\theta = k\phi + \phi_0, \quad (3.31)$$

where $k = 1$ and $\phi_0 = 0$ and $\pi/2$, respectively.

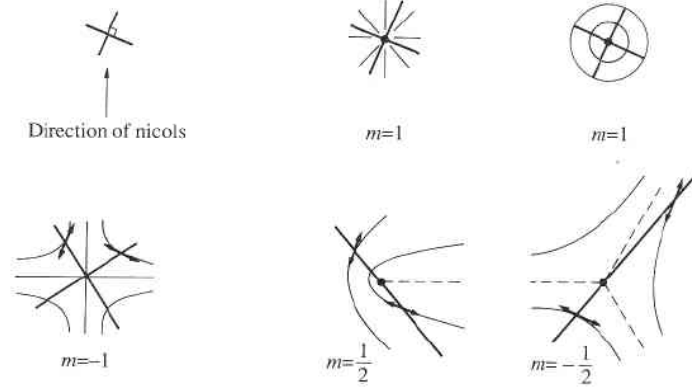


Figure 3.9: 2D nematic disclinations: $k = -1, -1/2, 1/2$, and two different $k = 1$ defects [15].

Within the one-constant approximation, the energy of defects can be calculated fairly straightforwardly. For 2D director field, the free energy density per unit volume reads

$$f_d = \frac{K}{2} (\nabla\theta)^2 \quad (3.32)$$

and at equilibrium $\nabla^2\theta = 0$. This equation is solved by Eq. (3.31). Recall that in cylindrical coordinates $\nabla f = \mathbf{e}_r \partial f / \partial r + (\mathbf{e}_\phi / r) \partial f / \partial \phi + \mathbf{e}_z \partial f / \partial z$ and $\nabla \cdot \mathbf{A} = (\partial / \partial r + 1/r) A_r + (1/r) \partial A_\phi / \partial \phi + \partial A_z / \partial z$. Thus the gradient of $\theta = k\phi + \phi_0$ is $\nabla\theta = (k/r)\mathbf{e}_\phi$ and the divergence of $\nabla\theta$ is 0. The energy per unit length of the defect line reads

$$\frac{F}{L} = \frac{K}{2} \int_a^{r_{max}} \left(\frac{k}{r}\right)^2 2\pi r dr = \pi K k^2 \int_a^{r_{max}} \frac{dr}{r} = \pi K k^2 \ln\left(\frac{r_{max}}{a}\right). \quad (3.33)$$

Here $a \sim 10$ nm is the radius of the defect core and r_{max} is the size of the region containing the defect. Note that the elastic energy of the defect diverges with r_{max} albeit logarithmically and that the total energy also includes the energy of the molten defect core.

The more important conclusion drawn from Eq. (3.33) is that the defect energy is proportional to the square of defect strength. From the topological perspective, a single $k = 1$ defect is equivalent to two $k = 1/2$ defects (Fig. 3.10) but its energy is bigger than twice the energy of a $k = 1/2$ defect pair:

$$\frac{F_{k=1}}{L} = \pi K \ln\left(\frac{r_{max}}{a}\right) \quad \text{vs.} \quad 2\frac{F_{k=1/2}}{L} = \frac{1}{2}\pi K \ln\left(\frac{r_{max}}{a}\right). \quad (3.34)$$

This means that only half-integer-strength defect lines are stable.

Figure 3.10: A $k = 1$ defect line can be split into two $k = 1/2$ defects.

In fact, line defects of integer strength of nematics in a capillary usually escape along the third dimension, leading to a director field of finite energy per unit length of $F/L \approx 3\pi K$ [15]. Adjacent regions where the director has escaped in opposite directions are separated by *point defects*. The description of point defects is much more complicated.

3.4 Nematic order parameter*

The director \mathbf{n} does not capture all aspects of nematic order — it only describes the preferred molecular orientation but not the magnitude of order or its biaxiality. The full nematic order parameter is defined based on the anisotropy of a suitable macroscopic quantity such as magnetic or electric susceptibility. In isotropic liquids, susceptibility is isotropic:

$$\mathbf{M} = \begin{bmatrix} \chi & & \\ & \chi & \\ & & \chi \end{bmatrix} \mathbf{H}. \quad (3.35)$$

On the other hand, in a uniaxial nematic with \mathbf{n} along z axis,

$$\mathbf{M} = \begin{bmatrix} \chi_{\perp} & & \\ & \chi_{\perp} & \\ & & \chi_{\parallel} \end{bmatrix} \mathbf{H} \quad (3.36)$$

where χ_{\perp} and χ_{\parallel} are susceptibilities perpendicular and parallel to the director, respectively. Thus the nematic order can be defined by the anisotropy of the magnetic susceptibility

$$Q_{ij} \propto \chi_{ij} - \frac{1}{3} \chi_{kk} \delta_{ij}. \quad (3.37)$$

The normalization is usually such that the largest eigenvalue is the degree of nematic order $S = \langle P_2(\cos \theta) \rangle$ [Eq. (3.13)]. Since Q_{ij} is proportional to the anisotropy of χ_{ij} its trace must vanish so that in a uniaxial nematic liquid crystal

$$\mathbf{Q} = \begin{bmatrix} -S/2 & & \\ & -S/2 & \\ & & S \end{bmatrix}. \quad (3.38)$$

In case of biaxial order, the eigenvalues corresponding to directions perpendicular to \mathbf{n} are not identical and

$$\mathbf{Q} = \begin{bmatrix} -S/2 - P & & \\ & -S/2 + P & \\ & & S \end{bmatrix}, \quad (3.39)$$

where P is the degree of biaxiality. As biaxial nematics are rather exotic, they will not be considered here; we merely note that P reads

$$P = \frac{1}{2} \langle \sin^2 \theta \cos 2\phi \rangle, \quad (3.40)$$

where ϕ is the azimuthal angle of the molecular long axis.

From the definition of the degree of order

$$S = \frac{1}{2} \langle 3 \cos^2 \theta - 1 \rangle \quad (3.41)$$

we see that if all molecules point along the director such that $\theta = 0$ for all molecules $S = 1$. This state corresponds to perfect nematic order (Fig. 3.11). If the molecules are oriented isotropically, $\langle \cos^2 \theta \rangle = 1/3$ and $S = 0$; this is the isotropic phase. A special kind of nematic orientational order is that with all molecules perpendicular to the director such that $\theta = \pi/2$. In this case $S = -1/2$.

In real nematics, S may reach 0.6 or so; on approaching the nematic-isotropic transition by heating, S decreases.

Figure 3.11: Nematic order with $S = 1$, isotropic order with $S = 0$, nematic order with $S = -1/2$.

3.5 Landau-de Gennes theory*

The tensorial order parameter \mathbf{Q} can be used to construct a Landau-type phenomenological theory of the nematic-isotropic transition. This theory is of interest both directly and because nematics are laden with topological defects where the director is not defined and the orientational order locally melts.

In the Landau-de Gennes theory, the free energy per unit volume is expanded in terms of the scalar invariants of \mathbf{Q} . Since \mathbf{Q} is traceless, there is no first-order invariant to rigid rotations. The second-order invariant is

$$Q_{ij}Q_{ji} \quad (3.42)$$

and the third-order invariant is

$$Q_{ij}Q_{jk}Q_{ki}. \quad (3.43)$$

The Landau-de Gennes expansion reads

$$f_{LdG} = f_0 + \frac{1}{2}a(T - T^*)Q_{ij}Q_{ji} - \frac{1}{3}bQ_{ij}Q_{jk}Q_{ki} + \frac{1}{4}c(Q_{ij}Q_{ji})^2. \quad (3.44)$$

a , b , and c are phenomenological constants and T^* is the so-called supercooling temperature.

This expansion is constructed such that both the isotropic phase with $S = 0$ and the nematic phase with $S > 0$ are possible solutions. The specific features of this theory are:

- i) At $T < T^*$, the coefficient of the second-order term is negative so that at low temperatures isotropic phase is unstable. This is why T^* is referred to as the supercooling temperature.
- ii) Since the states with $S = S' > 0$ and $S = -S' < 0$ are physically inequivalent, the third-order term is not forbidden. In order to favor states with positive S which are observed, $b < 0$.
- iii) In order to stabilize the theory and make the Landau expansion bounded from below, the fourth-order term is included.

The workings of the Landau-de Gennes theory is best seen by considering uniaxial homogeneous order with $P = 0$. This gives

$$f_{LdG} = f_0 + \frac{3a(T - T^*)}{4} S^2 - \frac{b}{4} S^3 + \frac{9c}{16} S^4. \quad (3.45)$$

The dependence of the free energy on S at a few representative temperatures is shown in Fig. 3.12.

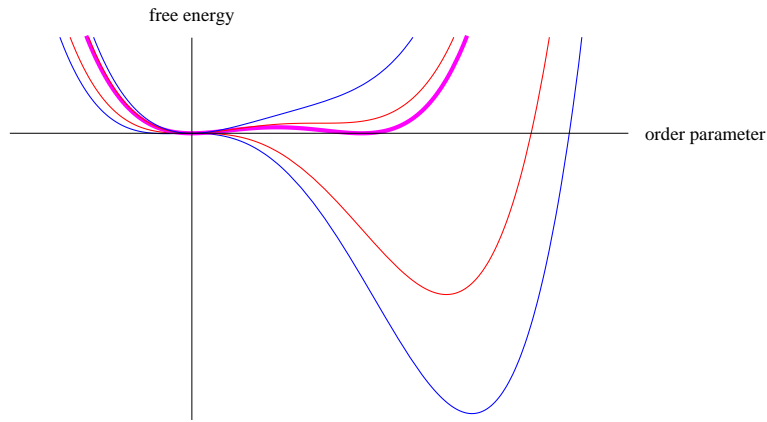


Figure 3.12: Landau-de Gennes theory of nematic-isotropic transition: Free energy at $T < T^*$ (bottom blue curve), $T = T_{NI}$ (thick magenta curve), and $T > T^{**}$ (top blue curve). The top and bottom red curves correspond to the supercooling temperature T^* and to the superheating temperature T^{**} , respectively.

Because of the third-order term, the nematic-isotropic transition is discontinuous and the degree of order S experiences a jump. Also calculated can be the phase transition temperature T_{NI} (where the free energies of the

isotropic and the nematic phase are the same)

$$T_{NI} = T^* + \frac{b^2}{27ac} \quad (3.46)$$

and the superheating temperature T^{**} of the nematic phase (corresponding to the state where the local minimum at $S > 0$ vanishes)

$$T^{**} = T^* + \frac{b^2}{24ac}. \quad (3.47)$$

The theory can be used to calculate all thermodynamic quantities of interest: Latent heat, critical exponents, susceptibility...

To describe the defects, the basic theory must be extended so as to account for inhomogeneous order. There are two elastic terms

$$f_{LaG}^{el} = \frac{1}{2}L_1 \frac{\partial Q_{ij}}{\partial x_k} \frac{\partial Q_{ij}}{\partial x_k} + \frac{1}{2}L_2 \frac{\partial Q_{ij}}{\partial x_j} \frac{\partial Q_{ik}}{\partial x_k}. \quad (3.48)$$

Note that the existence of just two terms is inconsistent with the Frank free energy which includes three elastic modes.

3.6 Smectic elasticity

Unlike simple fluids where the density is uniform, smectics are characterized by a periodically modulated density profile [16]. In a coordinate system where the z axis points normal to the layers,

$$\rho(\mathbf{r}) = \rho_0 + [\langle \psi \rangle \exp(iq_0 z) + c.c.]. \quad (3.49)$$

$\langle \psi \rangle$ is thus the smectic order parameter, its magnitude measuring the deviation of the density from the mean value ρ_0 . $q_0 = 2\pi/d$ is the wavevector corresponding to layer spacing d . $\langle \psi \rangle$ is a complex number and can be written as

$$\langle \psi \rangle = |\langle \psi \rangle| \exp(-iq_0 u(\mathbf{r})), \quad (3.50)$$

where $u(\mathbf{r})$ denotes the displacement of the layers from the equilibrium position. A constant u represents a translation of the whole smectic liquid crystal along the z axis, and this does not change the energy. Like in crystals, it is the spatial variation of the displacement that gives rise to elastic energy — the only difference being that smectics are 2D liquids and so it is only the displacement in the direction perpendicular to the layers that

matters. Like in crystals, we seek an elastic energy expanded in terms of ∇u in a manner consistent with the symmetry of smectics.

In smectics A, the molecules are arranged in layers and since they are perpendicular to the layers, the layer normal \mathbf{N} and the director \mathbf{n} coincide. This constraint is maintained by the low-energy splay director deformations, whereas the twist and bend deformations of the director field violate the constraint and are thus energetically more costly. This is illustrated in Fig. 3.13 which also shows that the splay deformation of the director field corresponds to the bend deformation of the layers at constant spacing, whereas the bend deformation of the director corresponds to compression/dilation of the layers [16].

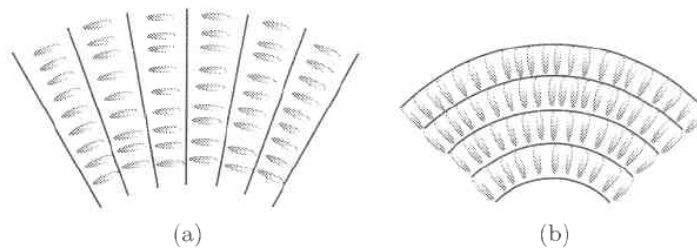


Figure 3.13: In smectics, the $\mathbf{N} = \mathbf{n}$ constraint is consistent with splay deformations but not with bend and twist deformations [16].

The layered structure of smectic implies that we need to consider separately the normal and the in-plane components of the gradient of u : $\nabla_z u$ changes the layer spacing whereas $\nabla_{\perp} u$ changes the orientation of the layers (Fig. 3.14).

Elastic theory

Since constant $\nabla_{\perp} u$ corresponds to a rigid rotation of the layers (and thus a different choice of the coordinate system), there is no elastic energy associated to it; it is only $\nabla_{\perp}^2 u$ (the radius of curvature) that gives rise to elasticity. We conclude that there must be two terms in the elastic free energy of smectic, one associated with layer compression and the other one to layer bending:

$$f_d = \frac{B}{2} (\nabla_z u)^2 + \frac{K_{11}}{2} (\nabla_{\perp}^2 u)^2. \quad (3.51)$$

B is the compression modulus and K_{11} is the bending constant — Fig. 3.13 suggests that the bending constant is the same as the nematic splay constant,

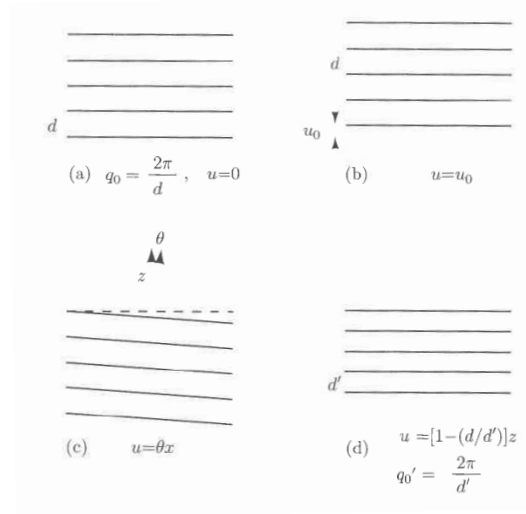


Figure 3.14: A modulated layer displacement along the layer normal changes layer spacing. The in-plane gradient of u tilts the layers and $\nabla_{\perp}^2 u$ bends them [16].

and we use the same symbol. This elastic energy is correct to lowest order, i.e., if \mathbf{N} and the director \mathbf{n} coincide.

It is clear from the dimensions of the two terms that the unit of B and K_{11} are not the same (J/m^3 and J/m , respectively). The relative weight of the two terms in Eq. (3.51) is conventionally parametrized by the penetration length (or depth)

$$\lambda = \sqrt{\frac{K_{11}}{B}} \quad (3.52)$$

so that

$$f_d = \frac{B}{2} \left[(\nabla_z u)^2 + \lambda^2 (\nabla_{\perp}^2 u)^2 \right]. \quad (3.53)$$

The penetration length is related to the length over which an imposed distortion of the smectic layers relaxes.

It can be shown that the deformation of smectic A in contact with an undulating hard surface decays exponentially with distance from the surface, the characteristic length being $l = 1/k^2 \lambda$ where k is the wave vector of the undulations. The shape of the wall at $z = 0$ is described by

$$z(x) = a \cos kx, \quad (3.54)$$

where k is the wavevector such that $ka \ll 1$, and we seek solutions of the form

$$u(x, z) = w(z) \cos kx. \quad (3.55)$$

Then the free energy density is

$$f_d = \frac{B}{2} \left(\frac{dw}{dz} \right)^2 \cos^2 kx + \frac{K_{11}}{2} k^4 w^2 \cos^2 kx \quad (3.56)$$

and the Euler-Lagrange equation for $w(z)$ reads

$$\frac{d^2 w}{dz^2} = \frac{K_{11}}{B} k^4 w = \lambda^2 k^4 w. \quad (3.57)$$

The solution is satisfying the boundary condition $w(z=0) = a$ is

$$w(z) = a \exp(-z/\ell), \quad (3.58)$$

where

$$\ell = \frac{1}{k^2 \lambda}. \quad (3.59)$$

Thus long-wavelength surface imperfections with small k give rise to large ℓ : Distortions of smectic order penetrate deep into the sample.

Elastic theory reexamined

The above theory of smectic elasticity has been derived based on the assumption that the director and the layer normal coincide $\mathbf{n} = \mathbf{N}$. This is not necessarily true and in a more general framework, these two vectors may depart from one another so that in a smectic where the undisturbed state is described by $\mathbf{n} = \mathbf{N} = \mathbf{e}_z$, deviations $\delta\mathbf{n}$ and $\delta\mathbf{N}$ need not be the same. This leads to an elastic energy of type

$$\frac{D}{2} (\delta\mathbf{n} - \delta\mathbf{N})^2. \quad (3.60)$$

So one can write the smectic elastic energy as a sum of layer compression term $B(\nabla_z u)^2$, the coupling term, and the Frank free energy penalizing the director distortions:

$$\begin{aligned} f_d = & \frac{B}{2} (\nabla_z u)^2 + \frac{D}{2} (\delta\mathbf{n} - \delta\mathbf{N})^2 \\ & + \frac{K_{11}}{2} (\nabla \cdot \mathbf{n})^2 + \frac{K_{22}}{2} (\mathbf{n} \cdot \nabla \times \mathbf{n})^2 + \frac{K_{33}}{2} (\mathbf{n} \times \nabla \times \mathbf{n})^2. \end{aligned} \quad (3.61)$$

Here $\mathbf{n} = \mathbf{e}_z + \delta\mathbf{n}$.

In terms of the layer displacement u , \mathbf{N} is to lowest order given by

$$\mathbf{N} = \mathbf{e}_z + \delta\mathbf{N} = \left(-\frac{\partial u}{\partial x}, -\frac{\partial u}{\partial y}, 1 \right). \quad (3.62)$$

The coupling term is minimal if $\delta\mathbf{n} = \delta\mathbf{N} = -\nabla_{\perp} u$ so that the splay Frank term reads

$$\nabla \cdot \delta\mathbf{n} = \nabla \cdot \left(-\frac{\partial u}{\partial x}, -\frac{\partial u}{\partial y}, 0 \right) = -\nabla_{\perp}^2 u. \quad (3.63)$$

So the splay term is the same as the layer bending term in Eq. (3.51). Yet another way of looking at it is to consider infinitesimal splay deformation of a director along \mathbf{e}_z : $\mathbf{n} = (ax, 0, 1)$ so that $\delta\mathbf{n} = ax\mathbf{e}_x$ and

$$u(x) = -\frac{ax^2}{2}. \quad (3.64)$$

The layer displacement does not depend on y or z and layer spacing is thus locally constant.

What about twist and bend deformations? The twist mode is given by, e.g.,

$$\delta\mathbf{n} = (ay, 0, 0) \quad (3.65)$$

so that if $\delta\mathbf{n}$ were equal to $\delta\mathbf{N} = -\nabla_{\perp} u$, the layer displacement would be given by

$$u(x, y) = -axy \quad (3.66)$$

and become infinitely large as $x \rightarrow \pm\infty$ even for very small separations along the twist direction y . Thus the condition $\mathbf{n} = \mathbf{N}$ cannot be met and twist deformation in smectics is not possible. [In fact, it is possible in the so-called twist grain boundary phases (Fig. 3.15) where the strain due to the above incompatibility is relaxed by screw dislocations.]

In a bend deformation,

$$\delta\mathbf{n} = (az, 0, 0) \quad (3.67)$$

and

$$u(x, z) = -axz \quad (3.68)$$

so that the layer displacement increases linearly with x , the slope being proportional to z . This fan-like pattern of layers does not preserve layer spacing and is thus impossible, so the coupling term cannot be zero. Thus the bend and the twist deformations are forbidden in smectics.

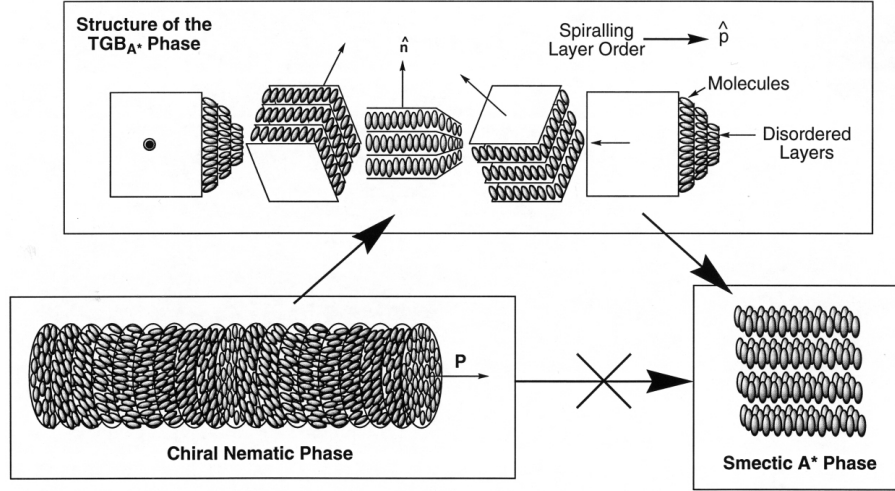


Figure 3.15: Twist grain boundary phase [10].

Nonlinear elasticity

The theory presented above is only invariant to infinitesimal rotations. For finite rotations, it does not account properly for the layer compression, which can be visualized as follows. If the layers are tilted from the z axis by an angle θ (which is approximately equal to $\approx \partial u / \partial x$ to lowest order), the equilibrium layer thickness d is decreased by

$$\delta d = d \sin\left(\frac{\pi}{2} - \theta\right) - d = d(\cos\theta - 1) \approx -d\frac{\theta^2}{2} \approx -d\frac{1}{2}\left(\frac{\partial u}{\partial x}\right)^2. \quad (3.69)$$

To maintain layer spacing, this decrease must be compensated by dilation in the z direction, the strain being given by $\partial u / \partial z$. So layer compression or dilation is really associated to the deviation of the gradients of u from

$$\frac{\partial u}{\partial z} - \frac{1}{2}\left(\frac{\partial u}{\partial x}\right)^2 \quad (3.70)$$

rather than by $\partial u / \partial z$ alone, and the corresponding energy density term is

$$\frac{B}{2} \left[\frac{\partial u}{\partial z} - \frac{1}{2}\left(\frac{\partial u}{\partial x}\right)^2 \right]^2. \quad (3.71)$$

In a mathematically complete way, a more complete free energy density valid for finite-angle rotations reads

$$f_d = \frac{B}{2} \left[\nabla_z u - \frac{1}{2} (\nabla u)^2 \right]^2 + \frac{K_{11}}{2} (\nabla_{\perp}^2 u)^2. \quad (3.72)$$

The nonlinear terms make the theory much more complicated.

Chapter 4

Polymers

Polymeric materials are based on long chain-like molecules. They are ubiquitous in everyday life as they are easy to mold and shape, and they are used for a range of products from textile fibers, plastic bags, and tires to glues and insulation materials such as styrofoam. Examples of synthetic polymers are polyvinylchloride and polyethylene; natural polymers include starch, DNA, etc. The basic building block of a polymer is a monomer and although the chemical composition of the monomer is important, most physical properties of polymers are generic and independent of chemistry. Instead they reflect the large length of polymers, their cross-linking (e.g., in gels), and entanglement.

The main property of any polymer is its degree of polymerization N , i.e., the number of monomers in the polymer. In most samples, the degree of polymerization varies from a chain to a chain and is polydisperse. In addition to linear polymers, there also exist ring and star polymers. Polymers may consist of a single type of monomer (homopolymers) or of two or more monomers (copolymers). Block copolymers consist of two typically incompatible covalently bonded blocks.

Polymers come in several states of matter. Polymer melts and solutions are far more viscous than simple liquids. Polymers readily form disordered glassy structures and crystalline polymers are usually not entirely crystallized; the crystalline spherulites are separated by a disordered liquid or a glassy matrix. Liquid-crystalline order in polymers may be due to elongated monomers included either as a main-chain or a side-chain component.

4.1 Single polymer chain*

Freely jointed chain*

The simplest model of the structure of a single linear polymer is the freely jointed chain where we assume that the bonds between the monomers of length a are perfectly flexible. If the repulsion between the monomers is neglected, the monomers can be considered the steps of a random walk in 3D. This means that the average end-to-end distance can be computed by averaging

$$\mathbf{r} = \mathbf{a}_1 + \mathbf{a}_2 + \dots + \mathbf{a}_N = \sum_{i=1}^N \mathbf{a}_i, \quad (4.1)$$

where \mathbf{a}_i is the vector describing the orientation of the i th monomer. The average end-to-end distance is a measure of polymer size:

$$\langle \mathbf{r}^2 \rangle = \langle \mathbf{r} \cdot \mathbf{r} \rangle = \left\langle \sum_{i,j=1}^N \mathbf{a}_i \cdot \mathbf{a}_j \right\rangle. \quad (4.2)$$

As the orientations of any two monomers are uncorrelated, it is instructive to write the double sum such that the cases with $i = j$ are separated from the rest:

$$\langle \mathbf{r}^2 \rangle = N \langle \mathbf{a}_i^2 \rangle + \left\langle \sum_{\substack{i,j=1 \\ i \neq j}}^N \mathbf{a}_i \cdot \mathbf{a}_j \right\rangle. \quad (4.3)$$

The second average is 0 and thus

$$\langle \mathbf{r}^2 \rangle = Na^2. \quad (4.4)$$

In other words, the size of the polymer coil is proportional to $N^{1/2}$. This is an example of a scaling law: This formula holds independent of the chemistry of a polymer. Chains where Eq. (4.4) is valid are referred to as *Gaussian* or *ideal chains*.

Radius of gyration

The size of the polymer coil can be measured in terms of $r = \sqrt{\mathbf{r}^2}$. Often it is expressed using the so-called *radius of gyration* defined by

$$r_G^2 = \frac{1}{N} \sum_{i=1}^N \langle |\mathbf{r}_i - \mathbf{r}_c|^2 \rangle, \quad (4.5)$$

where $\mathbf{r}_c = (1/N) \sum_{j=1}^N \mathbf{r}_j$ is the location of the center of mass. An equivalent definition of the radius of gyration is

$$r_G^2 = \frac{1}{2N^2} \sum_{i,j=1}^N \langle |\mathbf{r}_i - \mathbf{r}_j|^2 \rangle, \quad (4.6)$$

which can be proven easily. Note that

$$\sum_{i,j=1}^N (\mathbf{r}_i - \mathbf{r}_j)^2 = \sum_{i,j=1}^N [(\mathbf{r}_i - \mathbf{r}_c) - (\mathbf{r}_j - \mathbf{r}_c)]^2 \quad (4.7)$$

$$= \sum_{i,j=1}^N [(\mathbf{r}_i - \mathbf{r}_c)^2 - 2(\mathbf{r}_i - \mathbf{r}_c)(\mathbf{r}_j - \mathbf{r}_c) + (\mathbf{r}_j - \mathbf{r}_c)^2] \quad (4.8)$$

$$= N \sum_{i=1}^N (\mathbf{r}_i - \mathbf{r}_c)^2 - 2 \sum_{i=1}^N (\mathbf{r}_i - \mathbf{r}_c) \sum_{j=1}^N (\mathbf{r}_j - \mathbf{r}_c) + N \sum_{j=1}^N (\mathbf{r}_j - \mathbf{r}_c)^2 \quad (4.9)$$

Since $\sum_{i=1}^N (\mathbf{r}_i - \mathbf{r}_c) = 0$ we have $\sum_{i,j=1}^N (\mathbf{r}_i - \mathbf{r}_j)^2 = 2N \sum_{i=1}^N (\mathbf{r}_i - \mathbf{r}_c)^2$. — It can be shown that

$$r_G^2 = \frac{r^2}{6}. \quad (4.10)$$

Entropic elasticity*

The random-walk model can be used to compute the free energy of the freely jointed chain. To this end, consider the distribution of the end-to-end distances. Let us do this first in 1D; let a_x be the step length and N_+ and $N_- = N - N_+$ the numbers of steps to the right and to the left, respectively. The total end-to-end distance is

$$R_x = (N_+ - N_-)a_x \quad (4.11)$$

and the number of configurations giving a certain R_x is $\Omega_x = N!/(N_+!N_-!)$ so that $\ln \Omega_x$ can be approximated using Stirling formula:

$$\ln \Omega_x = N \ln N - N_+ \ln N_+ - (N - N_+) \ln(N - N_+) \quad (4.12)$$

$$= -N [f \ln f + (1 - f) \ln(1 - f)], \quad (4.13)$$

where $f = N_+/N$. $\ln \Omega_x$ peaks at $f = 1/2$ and since $d^2 \ln \Omega_x / df^2|_{f=1/2} = -4N$,

$$\ln \Omega_x = N \ln 2 - 2N \left(f - \frac{1}{2} \right)^2 = N \ln 2 - \frac{1}{2N} (N_+ - N_-)^2. \quad (4.14)$$

Now $N_+ - N_- = R_x/a_x$ so that

$$\Omega_x \propto \exp \left(-\frac{R_x^2}{2Na_x^2} \right). \quad (4.15)$$

In 3D, we also need to include Ω_y and Ω_z given by analogous expressions. As $R_x^2 + R_y^2 + R_z^2 = R^2$ and $a_x^2 = a_y^2 = a_z^2 = a^2/3$, we find that

$$\Omega \propto \exp \left(-\frac{3R^2}{2Na^2} \right) \quad (4.16)$$

and the complete 3D probability distribution is

$$P(\mathbf{r}, N) = \left(\frac{2\pi Na^2}{3} \right)^{-3/2} \exp \left(-\frac{3\mathbf{r}^2}{2Na^2} \right). \quad (4.17)$$

Using Boltzmann formula, we now have $S(\mathbf{r}, N) = k_B \ln P(\mathbf{r}, N)$ and $F = -TS$:

$$F(\mathbf{r}, N) = \frac{3k_B T \mathbf{r}^2}{2Na^2} + \text{const.} \quad (4.18)$$

Like in a Hookean spring, the free energy is proportional to extension squared. Note that the free energy is proportional to $k_B T$, which witnesses to the entropic origin of the elasticity.

Short-range correlations: Persistence length and Kuhn length

One may think that the freely jointed chain is a poor approximation and that the behavior of real chains where the bonds are not entirely free may change the above predictions a lot. In a more realistic model of a monomer-monomer bond, we assume that the i th and $(i+1)$ st monomer make an angle of θ and that the bond can rotate freely on the cone of angle θ . Then $\mathbf{a}_i = a\mathbf{u}_i$ where \mathbf{u} is the unit vector along \mathbf{a}_i ; \mathbf{v}_i and \mathbf{w}_i are the other two members of the orthonormal triad of basis vectors. Then

$$\mathbf{a}_{i+1} = a (\cos \theta \mathbf{u}_i + \sin \theta \cos \phi_i \mathbf{v}_i + \sin \theta \sin \phi_i \mathbf{w}_i), \quad (4.19)$$

where ϕ is the azimuthal angle describing the orientation of \mathbf{a}_{i+1} relative to \mathbf{a}_i . As the bond energy is independent of ϕ_i ,

$$\langle \mathbf{a}_i \cdot \mathbf{a}_{i+1} \rangle = a^2 \cos \theta. \quad (4.20)$$

Now

$$\mathbf{a}_{i+2} = a (\cos \theta \mathbf{u}_{i+1} + \sin \theta \cos \phi_{i+1} \mathbf{v}_{i+1} + \sin \theta \sin \phi_{i+1} \mathbf{w}_{i+1}) \quad (4.21)$$

$$= a \left(\cos^2 \theta \mathbf{u}_i + \cos \theta \sin \theta \cos \phi_i \mathbf{v}_i + \cos \theta \sin \theta \sin \phi_i \mathbf{w}_i \right. \\ \left. + \sin \theta \cos \phi_{i+1} \mathbf{v}_{i+1} + \sin \theta \sin \phi_{i+1} \mathbf{w}_{i+1} \right). \quad (4.22)$$

We compute the dot product $\mathbf{a}_i \cdot \mathbf{a}_{i+2}$. After averaging over ϕ_i and ϕ_{i+1} all terms but the first one vanish and we have

$$\langle \mathbf{a}_i \cdot \mathbf{a}_{i+2} \rangle = a^2 \cos^2 \theta. \quad (4.23)$$

We conclude that

$$\langle \mathbf{a}_i \cdot \mathbf{a}_{i+m} \rangle = a^2 \cos^m \theta. \quad (4.24)$$

As $\cos \theta < 1$, the orientational correlations between the monomers decay along the chain. The decay is characterized by the so-called *persistence length*

$$\ell_p = \lim_{N \rightarrow \infty} \langle \mathbf{u}_1 \cdot \mathbf{r} \rangle = \lim_{N \rightarrow \infty} \left\langle \mathbf{u}_1 \cdot a \sum_{i=1}^N \mathbf{u}_i \right\rangle \quad (4.25)$$

$$= a \sum_{m=0}^{\infty} \langle \mathbf{u}_1 \cdot \mathbf{u}_{1+m} \rangle = a \sum_{m=0}^{\infty} \cos^m \theta = \frac{a}{1 - \cos \theta}. \quad (4.26)$$

In a perfectly stiff chain $\theta = 0$ and $\ell_p \rightarrow \infty$. In general $\ell_p > a$.

After introducing the persistence length, we can visualize the polymer as a freely jointed chain of segments of length ℓ_p . The number of segments in the polymer is $N_s = N/g$ where g is the number of monomers in a segment. Thus we can reinterpret Eq. (4.4) to compute the end-to-end distance in the polymer:

$$\langle \mathbf{r}^2 \rangle = N_s \langle \mathbf{a}_s^2 \rangle = \frac{N}{g} \langle \mathbf{a}_s^2 \rangle, \quad (4.27)$$

where \mathbf{a}_s is the vector pointing from the beginning of a segment to the end. This result shows that the scaling law $\langle \mathbf{r}^2 \rangle \propto N$ is unchanged by short-range correlations.

The above illustration of the persistence length is based on the model where the angle between the neighboring monomers is fixed, which is not completely general. A more general definition is based on the bending energy of the polymer. If the polymer is approximated by a rod, then its bending energy per unit length is

$$\frac{F}{L} = \frac{EI}{2} \left(\frac{1}{R} \right)^2, \quad (4.28)$$

where E is Young modulus, I is the moment of inertia of the rod, and R and L are the radius of curvature and the length, respectively. The persistence length is defined as the length where the bending energy of a rod with $\ell_p = R$ (so that the rod is bent by an angle of 1 rd) is of the order of $k_B T$:

$$F = \frac{EI}{2} \left(\frac{1}{\ell_p} \right)^2 \ell_p = \frac{EI}{2\ell_p} \sim k_B T \quad (4.29)$$

so that

$$\ell_p = \frac{EI}{k_B T}. \quad (4.30)$$

A related measure of polymer stiffness is the *Kuhn length*. In a fully extended form, the length of a real chain is

$$r_{max} = N_s a_s, \quad (4.31)$$

where a_s is the segment length. Together with Eq. (4.27), this equation implies that the segment length can be defined by the ratio $\langle \mathbf{r}^2 \rangle / r_{max}$ and this is known as the Kuhn length:

$$\ell_K = \frac{\langle \mathbf{r}^2 \rangle}{r_{max}}. \quad (4.32)$$

The Kuhn and the persistence lengths are very similar.

Worm-like chain*

The freely-jointed chain is an idealization. A more realistic model of polymers is the *worm-like chain* (also referred to as the *Kratky-Porod model*) where the polymer is considered as a rod of finite flexibility. In this case, the orientational correlations along the chain decay exponentially

$$\langle \mathbf{t}(s=0) \cdot \mathbf{t}(s) \rangle = \exp \left(-\frac{s}{\ell_p} \right), \quad (4.33)$$

θ being the angle between tangents at s and in the origin. This can be seen by first considering the correlations between two adjacent monomers. If θ is the angle between them and the bending energy is given by

$$E_{bend} = \frac{EI}{2a} |\mathbf{u}_{i+1} - \mathbf{u}_i|^2 = -\frac{EI}{a} \mathbf{u}_{i+1} \cdot \mathbf{u}_i + const. = -\frac{EI}{a} \cos \theta + const. \quad (4.34)$$

(where EI is the bending rigidity of the polymer and a is the monomer length), then

$$\langle \mathbf{u}_i \cdot \mathbf{u}_{i+1} \rangle = \langle \cos \theta \rangle = \frac{\int \cos \theta \exp(\beta(EI/a) \cos \theta) \sin \theta d\theta}{\int \exp(\beta(EI/a) \cos \theta) \sin \theta d\theta} \quad (4.35)$$

$$= \coth \frac{\beta EI}{a} - \frac{a}{\beta EI} \approx 1 - \frac{a}{\beta EI}. \quad (4.36)$$

In the last step, we assumed that the temperature is low so that $\beta EI/a \gg 1$.

Persistence length is the distance beyond which the orientational correlations vanish, and short-range order is encoded by the exponential decay of correlations

$$\langle \mathbf{u}_i \cdot \mathbf{u}_j \rangle = \exp\left(-\frac{|i-j|a}{\ell_p}\right), \quad (4.37)$$

where a is the monomer size. For $j = i + 1$ we have

$$\langle \mathbf{u}_i \cdot \mathbf{u}_{i+1} \rangle = \exp\left(-\frac{a}{\ell_p}\right) \approx 1 - \frac{a}{\ell_p}. \quad (4.38)$$

By comparing this result to Eq. (4.36), we recover Eq. (4.30).

Now we calculate the end-to-end distance. In terms of the tangent vector \mathbf{t} along the polymer contour of length l ,

$$\langle \mathbf{r}^2 \rangle = \langle \mathbf{r} \cdot \mathbf{r} \rangle = \left\langle \int_0^l \mathbf{t}(s) ds \cdot \int_0^l \mathbf{t}(s') ds' \right\rangle \quad (4.39)$$

$$= \int_0^l \int_0^l \langle \mathbf{t}(s) \cdot \mathbf{t}(s') \rangle ds ds' \quad (4.40)$$

$$= \int_0^l ds \left[\int_0^s \exp\left(\frac{s'-s}{\ell_p}\right) ds' + \int_s^l \exp\left(-\frac{s'-s}{\ell_p}\right) ds' \right] \quad (4.41)$$

$$= \int_0^l ds \left\{ \ell_p \left[1 - \exp\left(-\frac{s}{\ell_p}\right) \right] + \ell_p \left[1 - \exp\left(-\frac{l-s}{\ell_p}\right) \right] \right\} \quad (4.42)$$

$$= 2\ell_p l \left\{ 1 - \frac{\ell_p}{l} \left[1 - \exp\left(-\frac{l}{\ell_p}\right) \right] \right\}. \quad (4.43)$$

The two limits of this result are: i) the freely jointed chain with $l \gg \ell_p$ and ii) the rigid rod with $l \ll \ell_p$. In the first case, the second term in the curly brackets can be neglected and we have $r^2 \propto l \propto N$ consistent with Eq. (4.4). In the second case the exponent can be expanded to second order $\exp(-l/\ell_p) \approx 1 - l/\ell_p + (l/\ell_p)^2/2 + \dots$ and the curly bracket reduces to $l/(2\ell_p)$ so that $r^2 = l^2$ as expected.

Steric repulsion between monomers: Expanded chains*

The restricted bond angle model represents a simple way of incorporating short-range correlations in the description of the polymer. The correlations between the more distant monomers due to the finite size of the monomers are far more important than the short-range correlations, and they are most straightforwardly accounted for by Flory's model. In this model, the polymer is treated as a gas of monomers contained in a box of volume $\sim r^3$. The free volume available to the monomers is

$$V_{free} = r^3 - NV_{excl}, \quad (4.44)$$

where $V_{excl} = V_1/2$ is the excluded volume per monomer and V_1 is the volume of one monomer; $1/2$ accounts for the pairwise nature of steric repulsion. Like in Onsager theory of nematic order, in the limit of small density the excess free energy due to excluded-volume interactions is

$$F^{ex} = Nk_B T \frac{NV_1}{2r^3}. \quad (4.45)$$

This term favors as large a size of polymer as possible and is opposed by the elasticity of the chain. The total free energy thus reads

$$F = k_B T \frac{3r^2}{2Na^2} + k_B T \frac{N^2 V_1}{2r^3}. \quad (4.46)$$

By minimizing this expression with respect to r we obtain $3r/(Na^2) - 3V_1 N^2/(2r^4) = 0$. Since $V_1 \propto a^3$ we find that in equilibrium the size of an expanded chain is

$$r \propto aN^{3/5}. \quad (4.47)$$

The exponent of $3/5$ is very close to the exact result of the renormalization-group analysis which gives $r \propto aN^{0.588}$. The thus defined size of the coil is sometimes referred to as the *Flory radius*.

4.2 Coil-globule transition

The above discussion of polymer coil size relies exclusively on the entropic interactions of the neighboring and distant parts of chain. But the structure of the polymer chain also depends on its interaction with the solvent. If the polymer and the solvent repel each other, polymer coil collapses. On the other hand, if the polymer-solvent contacts are favorable, the coil remains swollen.

The energy of interaction between the monomers themselves as well as between the monomers and the solvent can be included in terms of a simple mean-field theory based on local concentrations of the monomers and the solvent within the volume occupied by the coil. The local density of monomers within the coil is

$$\rho_l \sim \frac{N}{r^3}. \quad (4.48)$$

Then the number of monomer-monomer contacts within the coil volume is

$$N_{pp} = \frac{1}{2}zNV_1\rho_l, \quad (4.49)$$

where V_1 is the monomer volume. This result can be interpreted as follows: The total volume occupied by the monomers of a chain is NV_1 , and the number of times a monomer will be located within this volume (and thus be next to another monomer) is given by the product of NV_1 and ρ_l . The number of monomer-monomer contacts is then $NV_1\rho_l$ multiplied by z , the number of neighbors of each site occupied either by a monomer or a solvent molecule. The final result must be divided by 2 so as not to count each monomer-monomer pair twice.

In a similar way, we calculate the number of monomer-solvent contacts which is

$$N_{ps} = zNV_1(\rho_0 - \rho_l). \quad (4.50)$$

Here $\rho_0 - \rho_l$ is the local density of the solvent; we denote the overall density by ρ_0 . Thus $NV_1(\rho_0 - \rho_l)$ is the number of times a solvent molecule is located within the total volume occupied by the monomers, thereby establishing a monomer-solvent contact. Again the number of contacts depends on the coordination number z . — The number of actual solvent-solvent contacts is the reference number of solvent-solvent contacts N_{ss}^0 decreased by N_{pp} and N_{ps} :

$$N_{ss} = N_{ss}^0 - N_{pp} - N_{ps} = N_{ss}^0 - \frac{1}{2}zNV_1\rho_l - zNV_1(\rho_0 - \rho_l). \quad (4.51)$$

We assume that the interactions are pairwise additive, the monomer-monomer, monomer-solvent, and solvent-solvent interaction energies being ϵ_{pp} , ϵ_{ps} , and ϵ_{ss} , respectively. The total interaction energy is then

$$U^{int} = \epsilon_{pp}N_{pp} + \epsilon_{ps}N_{ps} + \epsilon_{ss}N_{ss} \quad (4.52)$$

$$= (\epsilon_{pp} - \epsilon_{ss})\frac{1}{2}zNV_1\rho_l + (\epsilon_{ps} - \epsilon_{ss})zNV_1(\rho_0 - \rho_l) + \epsilon_{ss}N_{ss}^0 \quad (4.53)$$

$$= (\epsilon_{pp} + \epsilon_{ss} - 2\epsilon_{ps})\frac{1}{2}zNV_1\rho_l + (\epsilon_{ps} - \epsilon_{ss})zNV_1\rho_0 + \epsilon_{ss}N_{ss}^0 \quad (4.54)$$

$$= (\epsilon_{pp} + \epsilon_{ss} - 2\epsilon_{ps})\frac{1}{2}zNV_1\rho_l + const. \quad (4.55)$$

In the final form, we only retain the term that depends on the density (and thus on the structure) of the polymer coil. Now we introduce the dimensionless *interaction parameter* χ defined by

$$\chi k_B T = \frac{1}{2}z(2\epsilon_{ps} - \epsilon_{pp} - \epsilon_{ss}) \quad (4.56)$$

If interaction of monomer and solvent is favorable, i.e., $\epsilon_{ps} < (\epsilon_{pp} + \epsilon_{ss})/2$, then $\chi < 0$.

In terms of χ ,

$$U^{int} = -Nk_B TV_1\chi\rho_l + const. \quad (4.57)$$

This result can now be used to renormalize the excluded-volume entropic free energy of the chain [Eq. (4.45)] which has the same functional form. Recall that $\rho_l \sim N/r^3$ and

$$F^{ex} + U^{int} = k_B T (1 - 2\chi) \frac{N^2 V_1}{2r^3} + const. \quad (4.58)$$

We see that $F^{ex} + U^{int}$ is positive if $\chi < 1/2$. In this case, the $F^{ex} + U^{int}$ plays the same role as F^{ex} alone and the qualitative behavior of the chain is the same as if there were no monomer-monomer, monomer-solvent, and solvent-solvent interactions. The size of the polymer coil is given by Eq. (4.47). This is the so-called *good solvent behavior*.

On the other hand, if $\chi = 1/2$ the $F^{ex} + U^{int} = 0$ and the effect of excluded volume is counterbalanced by the attraction between the monomers and the solvent molecules. This is the *theta condition*. If $\chi > 1/2$, the incompatibility of the solvent and the polymer is so large that the total energy

$$F = k_B T \frac{3r^2}{Na^2} + k_B T (1 - 2\chi) \frac{N^2 V_1}{2r^3}. \quad (4.59)$$

is dominated by the second term which is negative such that the equilibrium size of the chain tends to 0 and the polymer forms a *globule*.

4.3 Polymer solutions

Dilute, semidilute, and concentrated solutions

Knowing how isolated polymer chains behave describes the *dilute* polymer solutions. Due to their expanded rather than compact size, it takes a very small concentration of polymers to enter the so-called *semidilute* regime where the polymers partly overlap (Fig. 4.1)

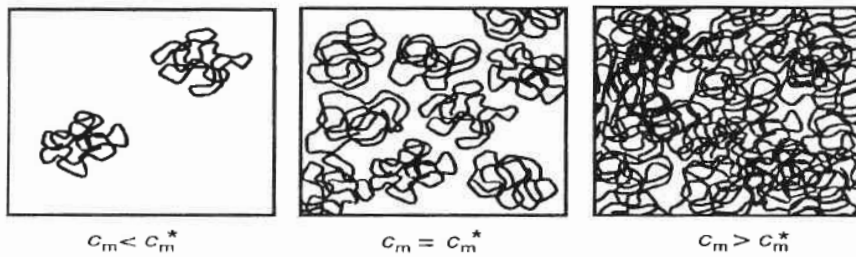


Figure 4.1: Dilute, semidilute, and concentrated polymer solutions [2].

Let us compute the critical density of monomers needed for the overlap of coils. We define it by

$$\rho^* = \frac{N}{r^3}. \quad (4.60)$$

Since $r = aN^{3/5}$ [Eq. (4.47)] we have

$$\rho^* = \frac{N}{a^3 N^{9/5}} = \frac{1}{a^3 N^{4/5}}. \quad (4.61)$$

In terms of monomer volume V_1 , the packing fraction is given by

$$\eta^* = V_1 \rho^* = \frac{V_1}{a^3 N^{4/5}} \quad (4.62)$$

and since V_1 is typically smaller than a^3 (this was not taken into account in the above discussion of the Flory theory) we conclude that

$$\eta^* < N^{-4/5}. \quad (4.63)$$

For a polymer of $N \sim 10^4$ the critical packing fraction is very small, less than $\eta^* \sim 10^{-16/5} \approx 0.001$. Polymer solutions are dilute only if $\eta < \eta^*$ whereas for η between 0.001 and 0.1 the interpenetration of chains is considerable.

Osmotic pressure

At low concentrations, the osmotic pressure of a polymer solution can be described by the virial expansion [cf. Eq. (2.76)]

$$\Pi = k_b T (B_1 \rho + B_2 \rho^2 + \dots). \quad (4.64)$$

This equation differs from the ordinary low-molar-mass virial equation of state [Eq. (2.76)] in that B_1 is not 1 but

$$B_1 = \frac{1}{N}. \quad (4.65)$$

To understand this, note that the osmotic pressure is due to translational motion of polymers rather than to the motion of chains. Thus the pressure does not depend on the monomer density ρ but on the polymer density

$$\rho_p = \frac{\rho}{N}. \quad (4.66)$$

Thus

$$\beta \Pi = \frac{\rho}{N} + B_2 \rho^2 + \dots = \frac{\rho}{N} (1 + N B_2 \rho + \dots) \quad (4.67)$$

and

$$\frac{\beta \Pi}{\rho_p} = 1 + N B_2 \rho + \dots \quad (4.68)$$

If the experimentally measured osmotic pressure multiplied by β/ρ_p is plotted against the so-called overlap ratio

$$x = \frac{\rho}{\rho^*} = \rho_p r^3 \quad (4.69)$$

the measurements for all polymers collapse onto a single curve as shown in Fig. 4.2 [2]. This suggests that the right-hand side of Eq. (4.68) depends on x . In the dilute limit, it must be equal to 1 and we may expect that it can be expanded in a power series of x . By keeping only the linear term we have

$$\frac{\beta \Pi}{\rho_p} = 1 + A_2 x + \dots \quad (4.70)$$

where A_2 must be a constant. By comparing Eq. (4.68) and Eq. (4.70) $N B_2 \rho = A_2 \rho_p r^3$ and

$$B_2 = A_2 \frac{\rho_p r^3}{N \rho} = A_2 \frac{r^3}{N^2}. \quad (4.71)$$

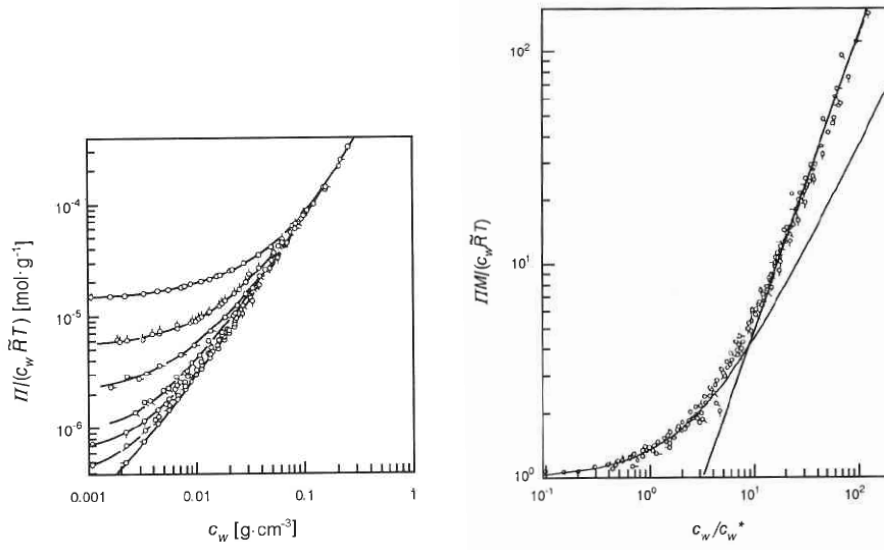


Figure 4.2: Osmotic pressure in poly(α -methylstyrene) in toluene plotted against mass density (a) for a range of degrees of polymerization spanning 2 orders of magnitude. Panel b shows the same curves multiplied by β/ρ_p and plotted against the overlap ratio x [2].

By inserting this result in Eq. (4.68) we find that

$$\frac{\beta\Pi}{\rho_p} = 1 + A_2 r^3 \rho_p + \dots \quad (4.72)$$

which suggests that the dilute solution of coils behaves as a gas of hard spheres of density ρ_p and size $A_2 r^3$ if we interpret the coefficient in front of ρ_p as the hard-core second virial coefficient [Eq. (2.86)]. This behavior is understandable as the overlap of two polymer coils involves many monomer-monomer contacts.

The regime of entangled semi-dilute solution $x > 1$ is qualitatively very different. Here the coils interpenetrate each other very much and pressure should not depend on the degree of polymerization since the entanglement happens at the level of segments much smaller than the chain length N . This means that the right-hand side in Eq. (4.67) is independent of N . This is only possible if the bracket, which depends on x , is a suitable power of x . Thus

$$\frac{\rho}{N} x^k \quad (4.73)$$

must be independent of N . Since $x = \rho_p r^3 = \rho r^3/N$ and $r \sim N^{3/5}$ [Eqs. (4.69) and (4.47)] $\sim \rho N^{9/5}/N = \rho N^{4/5}$, this means that

$$N^{4k/5-1} \quad (4.74)$$

is independent of N which implies that $k = 5/4$. We now use this result to derive the dependence of the osmotic pressure on density in the semi-dilute regime. From Eqs. (4.73) and Eqs. (4.69) it follows that $\beta\Pi \sim \rho \times \rho^{5/4} = \rho^{9/4}$ or

$$\frac{\beta\Pi}{\rho} \sim \rho^{5/4}. \quad (4.75)$$

This result is consistent with experiments.

4.4 Dynamical models

After the force acting on the ends of a polymer is turned off, the polymer shape recoils with a characteristic relaxation time. In an unentangled melt, this relaxation can be described by the so-called *Rouse modes*. The main process of interest in an entangled melt is creep taking place after a terminal time when the polymer begins to flow due to the applied shear stress. The simplest model accounting for the observed behavior is the *reptation theory*.

Rouse modes

The Rouse picture of relaxation in polymers relies on a bead-and-spring representation of the polymer where the chain is divided into the *Rouse sequences* represented by beads; the bonds between the sequences are represented by springs of spring constant b_R . The effect of neighboring chains is subsumed in an effective viscosity experienced by each monomer of the chain in question. In the overdamped limit, the equation of motion of bead j reads

$$\zeta_R \frac{d\mathbf{r}_j}{dt} = b_R (\mathbf{r}_{j+1} - \mathbf{r}_j) + b_R (\mathbf{r}_{j-1} - \mathbf{r}_j), \quad (4.76)$$

the two terms on the right hand side corresponding to the two springs acting on bead j . ζ_R is the friction coefficient. The spring constants b_R can be estimated using Eq. (4.18) which suggests that

$$b_R = \frac{3k_B T}{a_R^2}, \quad (4.77)$$

where a_R is the length of the Rouse sequence. The x, y , and z directions are decoupled and equivalent so that we need to consider only one of them:

$$\zeta_R \frac{dz_j}{dt} = b_R (z_{j+1} - z_j) + b_R (z_{j-1} - z_j) \quad (4.78)$$

and the solutions are of the form

$$z_j \sim \exp\left(-\frac{t}{\tau}\right) \exp(ij\delta). \quad (4.79)$$

The time dependence is exponential, the characteristic time being τ , whereas the spatial dependence along the chain is wave-like and δ represents the phase shift of neighboring beads. From Eq. (4.78) it follows that

$$\tau^{-1} = \frac{b_R}{\zeta_R} (2 - 2 \cos \delta) = \frac{4b_R}{\zeta_R} \sin^2 \frac{\delta}{2}. \quad (4.80)$$

The relaxation rate τ^{-1} is smallest at small δ which corresponds to long-wavelength modes, and maximal at $\delta = \pm\pi$ which corresponds to an alternating displacement pattern (Fig. 4.3).

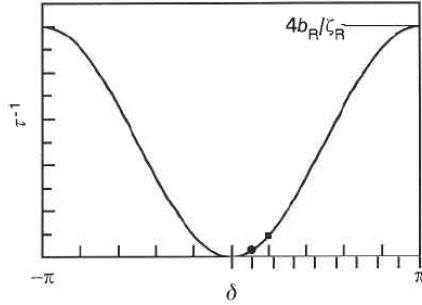


Figure 4.3: Dependence of the Rouse relaxation rate on the phase shift δ [2].

Let us now determine the slowest relaxation rate. The phase shifts are determined by the boundary conditions. At the two ends of the chain, the tensile force vanishes and thus the displacement of the first and second bead must be identical, and the same holds for the next-to-last and last bead:

$$z_1 - z_0 = 0 \quad \text{and} \quad z_{N_R-1} - z_{N_R-2} = 0. \quad (4.81)$$

In the continuous limit

$$\left. \frac{dz}{dj} \right|_{j=0} = 0 \quad \text{and} \quad \left. \frac{dz}{dj} \right|_{j=N_R-1} = 0. \quad (4.82)$$

The solutions of Eq. (4.78) compatible with these boundary conditions are

$$z_j \sim \exp\left(-\frac{t}{\tau}\right) \cos(ij\delta) \quad (4.83)$$

if δ satisfies

$$\sin((N_R - 1)\delta) = 0, \quad (4.84)$$

which follows from the second boundary condition (Fig. 4.4).

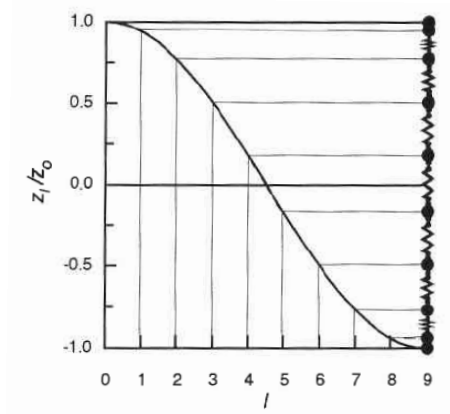


Figure 4.4: Displacement pattern of the $m = 1$ Rouse mode [2].

The allowed phase shifts are given by

$$\delta = \frac{m\pi}{N_R - 1} \quad (4.85)$$

where $m = 0, 1, \dots, N_R - 1$. The phase shift δ and thus the relaxation rate $\tau^{-1}(\delta)$ is smallest in the lowest-order Rouse mode with $m = 1$. This rate is usually referred to as the Rouse rate and is given by

$$\tau_R^{-1} = \frac{4b_R}{\zeta_R} \sin^2 \frac{\pi}{2(N_R - 1)} \approx \frac{b_R}{\zeta_R} \frac{\pi^2}{(N_R - 1)^2}, \quad (4.86)$$

where $\sin(\delta/2)$ was approximated by $\delta/2$. Using Eq. (4.77), we find that

$$\tau_R^{-1} = \frac{3\pi^2 k_B T}{\zeta_R a_R^2 (N_R - 1)^2}. \quad (4.87)$$

It seems that this result depends on the choice of the size of Rouse sequences a_R but note that $r^2 = a_R^2 (N_R - 1)$. Thus

$$\tau_R = \frac{\zeta_R / a_R^2}{3\pi^2 k_B T} r^4. \quad (4.88)$$

This result makes sense only if the ratio ζ_R/a_R^2 is independent of the choice of the Rouse sequences, which is true if ζ_R is proportional to the number of monomers in the sequence $N/(N_R - 1)$. In this case

$$\tau_R \sim N^2, \quad (4.89)$$

which is consistent with experimental results.

Reptation theory*

Polymer melts flow but their response to shear stress is markedly viscoelastic. This is due to the entanglement of the polymer chains, and the characteristic time after which the melt flows depends on the degree of polymerization. The typical time dependence of the shear modulus is shown in Fig. 4.5. At short times, the shear modulus is constant but after the so-called *terminal time* τ_T it decreases dramatically and the polymer behaves as a liquid.

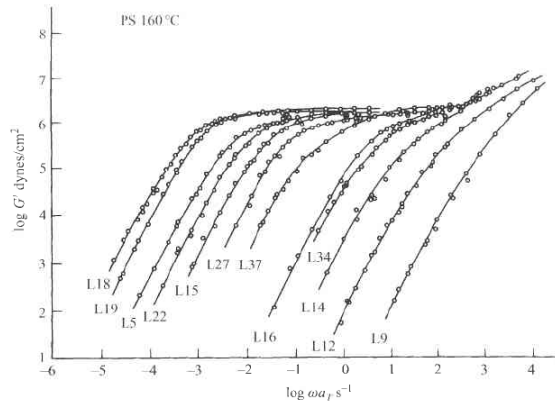


Figure 4.5: Storage modulus in polystyrene vs. frequency for a range of degrees of polymerization (which decreases from the leftmost curve to the rightmost curve). The small-frequency edge of the plateau corresponds to the terminal time. [7].

Experiments show that the terminal time increases with the degree of polymerization approximately as $\tau_T \propto N^{3.4}$. This dependence can be accounted for by the simple tube model of polymer motion.

The main idea behind the reptation theory is that the entangled chains cannot cross each other and that shear stress can only cause them to move

along the tubes in which they reside. Lateral motion is not possible and this kind of motion is referred to as *reptation* (Fig. 4.6).

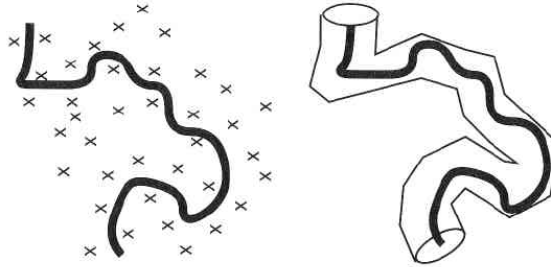


Figure 4.6: Reptation: The neighboring polymers form a tube in which the chain moves, and the time needed to leave the tube is the terminal time [1].

In this picture, a chain has relaxed after it has moved out of the tube of its own length. The time needed for this lengthwise motion is the terminal time, and it can be calculated by assuming that the motion is overdamped so that the resistance force is proportional to the velocity, the ratio being the mobility per monomer μ_m . Moving an N -monomer chain at a given speed takes an N -times larger force so the chain mobility is

$$\mu_{chain} = \frac{\mu_m}{N}. \quad (4.90)$$

The diffusion coefficient characterizing the 1D Brownian motion of the chain within the tube can be related to the mobility by the Einstein formula

$$D_{chain} = k_B T \mu_{chain} = \frac{k_B T \mu_m}{N}. \quad (4.91)$$

The chain escapes from the tube after it traverses a distance $L = Nl_m$ where l_m is the monomer length. For 1D Brownian motion, $r^2 = Dt$ and so

$$\tau_T = \frac{L^2}{D_{chain}} = \frac{N^2 l_m^2}{k_B T \mu_m / N} \propto N^3. \quad (4.92)$$

This result is reasonably close to the experimentally observed relation $\tau_T \sim N^{3.4}$. This basic version of the reptation theory can be refined by accounting for the reptating motion of the neighboring chains and for the fluctuations of the chain length — the length of chain undergoing Brownian motion within the tube fluctuates in time, its average being proportional to \sqrt{N} .

4.5 Gels*

The polymers in a solution can be connected by cross-links so as to form a macroscopic network of a finite shear modulus. This state is referred to as a *gel*. The cross-links may be covalent bonds (in the *chemical gels* such as rubber which is formed by vulcanization where linear polyisoprene chains of natural rubber are cross-linked by sulphur) or physical (in the *physical gels* such as gelatin). In physical gels, gelation can be controlled by temperature and physical gels can often be melted into sols (i.e., a colloidal suspension of solid particles forming the gel) by heating. Conversely, gelation can be induced by cooling.

During the sol-gel transition, the number of cross-links increases continuously but the transition is abrupt: As soon as the number of cross-links exceeds a threshold, the gel is formed. Gelation is an example of a percolation transition best illustrated by a lattice model (Fig. 4.7). As the number of bonds is increased, two processes take place: i) the number of clusters of connected sites increases, and ii) the average size of clusters grows. At a certain threshold, the biggest cluster spans the whole lattice. Thus the gel state is reached.

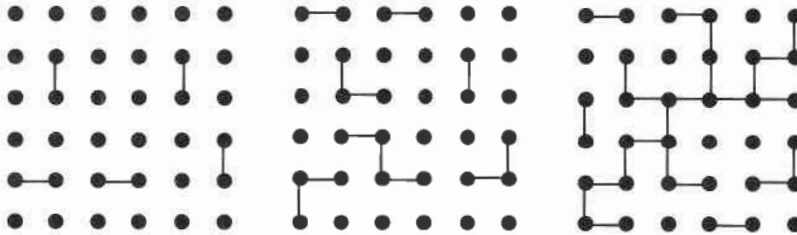


Figure 4.7: Lattice model of percolation [1].

Flory-Stockmayer theory of gelation*

The simplest theory of gelation is due to Flory and Stockmayer. Consider the Cayley tree, a radial network originating in a vertex connected to z nearest neighbors such that each of them branch to $z - 1$ second-generation vertices (Fig. 4.8). The vertices represent monomers and the bonds represent the cross-links. Denote the probability that a bond on the Cayley tree exists by f . If all bonds are independent, then on average each vertex of the n th

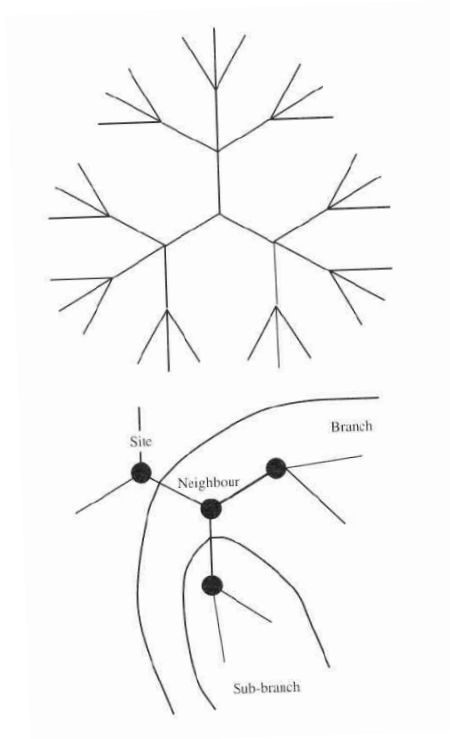


Figure 4.8: Cayley tree of $z = 4$ (top, [1]). A detail of the $z = 3$ Cayley tree (bottom, [1]).

generation is connected to $f(z - 1)$ vertices of the $(n + 1)$ st generation [19]. If the average number of bonds connecting a vertex to its outer neighbors is smaller than 1, then the number of different outward paths decreases in each generation. As $n \rightarrow \infty$, the probability of finding a connected sequence of bonds vanishes. Thus the *percolation threshold* specifying the minimal bond probability needed for the existence of an infinite cluster reads

$$f_c > \frac{1}{z - 1}. \quad (4.93)$$

Below the threshold $f < f_c$, the cluster is finite and this corresponds to the sol. Above the threshold cluster size is infinite and this is the gel. — The above model of gelation is rather abstract but it correctly captures the main features of gels.

Beyond the threshold, there exists one infinitely large cluster but not all vertices are a part of it. To quantify the degree of connectedness of the gel,

we now estimate the fraction of bonds in the cluster; this is referred to as the *gel fraction*.

The probability that a given vertex is connected to the infinite cluster P is related to the probability that the vertex is not connected to it by a given bond emanating from it Q . There are a total of z bonds emanating from the vertex and the probability that the vertex is connected to one of its neighbors but not connected via this neighbor to the infinite cluster is fQ^{z-1} (the power of $z - 1$ is due to the fact that one of the bonds of the neighbor is connected to the vertex and the remaining $z - 1$ bonds do not lead to the infinite cluster).

Now we can spell out two possible scenarios why the vertex may not be connected to the infinite cluster via a given neighbor: i) because it is connected to a neighbor but that neighbor is not connected to the infinite cluster (probability fQ^{z-1}) or ii) because it is not connected to this neighbor (probability $1 - f$). Thus

$$Q = 1 - f + fQ^{z-1}. \quad (4.94)$$

In the $z = 3$ Cayley tree, this is a quadratic equation and the solutions are $Q = 1$ and $Q = (1 - f)/f$.

The probability that a vertex is not connected to the infinite cluster is Q^z and the probability that it is connected to a given neighbor by a given bond but not to infinity is fQ^z . This probability is equal to the probability that this bond exists f less the probability that the vertex is connected to the infinite cluster P . Thus

$$f - P = fQ^z \quad (4.95)$$

and

$$P = f - fQ^z = f(1 - Q^z). \quad (4.96)$$

For $z = 3$, we find that for $f > f_c$ where $Q = (1 - f)/f$ the probability that a vertex is a part of the infinite cluster relative to the fraction of bonds formed is

$$\frac{P}{f} = 1 - \left(\frac{1-f}{f}\right)^3. \quad (4.97)$$

For $f < f_c$ where $Q = 1$ $P/f = 0$. Figure 4.9 shows that the gel fraction grows rapidly as $f > f_c$, approaching 1 as $f \rightarrow 1$.

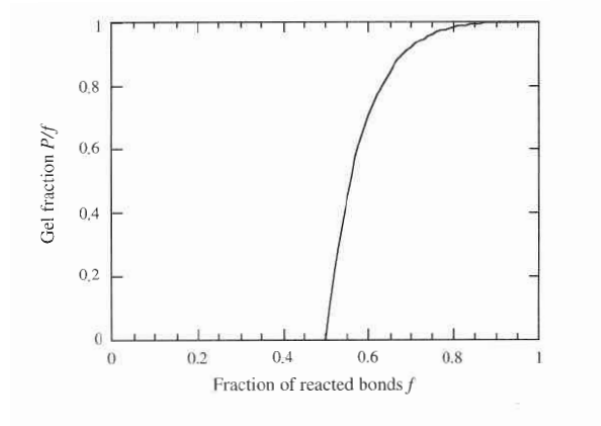


Figure 4.9: Gel fraction vs. the fraction of bonds formed for $z = 3$ [1].

Rubber elasticity

Rubber is a polymer melt with many random cross-links such that the chains form a macroscopic network. Locally rubber is liquid but on length scales larger than the average separation of cross-links it has a finite shear modulus. The classical theory of rubber elasticity rests upon the assumption of affine deformation of the material whereby the cuboid of reference dimensions L_x, L_y , and L_z is deformed such that its edges after the deformation are $\lambda_x L_x, \lambda_y L_y$, and $\lambda_z L_z$. Thus each point at (x, y, z) moves to $(\lambda_x x, \lambda_y y, \lambda_z z)$. The affine character of this transformation is ensured by the cross-links which prevent a shear-free, liquid-like response of the material.

The elastic free energy can be calculated by considering a chain connecting two cross-links. If one of them is chosen as the origin and the other one is at (x, y, z) , then the reference end-to-end distance is $r' = \sqrt{x^2 + y^2 + z^2}$. After deformation, the end-to-end distance is changed and it reads $r = \sqrt{\lambda_x^2 x^2 + \lambda_y^2 y^2 + \lambda_z^2 z^2}$. The corresponding change of free energy within the freely-jointed-chain model [Eq. (4.18)] is

$$F_{chain} = \frac{3k_B T}{2Na^2} [(\lambda_x^2 - 1)x^2 + (\lambda_y^2 - 1)y^2 + (\lambda_z^2 - 1)z^2]. \quad (4.98)$$

As the cross-links were introduced in the polymer melt in a stress-free state, $\langle x^2 \rangle = \langle y^2 \rangle = \langle z^2 \rangle = Na^2/3$ if we assume that the number of monomers between all adjacent cross-links along the chain is the same (which is admittedly a rough approximation); the reference state is assumed isotropic.

Thus

$$\langle F_{chain} \rangle = \frac{k_B T}{2} (\lambda_x^2 + \lambda_y^2 + \lambda_z^2 - 3). \quad (4.99)$$

To obtain the free energy per unit volume, this result is now multiplied by the number of chains between cross-links per unit volume ρ_p . For a uniaxial constant-volume deformation, $\lambda_z = \lambda$ and $\lambda_x = \lambda_y = 1/\sqrt{\lambda}$ (so that the total volume of the cuboid $\lambda_x L_x \times \lambda_y L_y \times \lambda_z L_z = L_x L_y L_z$ is conserved) giving the free energy density of the ideal rubber

$$f_{rubber} = \rho_p \langle F_{chain} \rangle = \frac{\rho_p k_B T}{2} \left(\lambda^2 + \frac{2}{\lambda} - 3 \right). \quad (4.100)$$

The unstressed state corresponds to $\lambda = 1$ and so the strain is given by $u_{zz} = \lambda - 1$. The stress can be calculated using the constitutive law

$$p_{ik} = \frac{\partial f_{rubber}}{\partial u_{ik}} \quad (4.101)$$

which implies that

$$p_{zz} = \frac{\partial f_{rubber}}{\partial u_{zz}} = \frac{\partial f_{rubber}}{\partial \lambda} = \rho_p k_B T \left(\lambda - \frac{1}{\lambda^2} \right). \quad (4.102)$$

This non-Hookean relation is obeyed by real rubbers for λ between about 50% and 150%, i.e. for strains up to 50% (Fig. 4.10).

For small deformations, Eq. (4.102) can be linearized:

$$p_{zz} = \rho_p k_B T \left[(1 + u_{zz}) - \frac{1}{(1 + u_{zz})^2} \right] \quad (4.103)$$

$$\approx \rho_p k_B T [1 + u_{zz} - (1 - 2u_{zz})] \quad (4.104)$$

$$\approx 3\rho_p k_B T u_{zz} \quad (4.105)$$

so that the Young modulus is

$$E = 3\rho_p k_B T \quad (4.106)$$

and the shear modulus is

$$\mu = \rho_p k_B T. \quad (4.107)$$

[Recall that $E = 2\mu(1 + \sigma)$ such that for an incompressible material where $\sigma = 1/2$ $E = 3\mu$.]

To appreciate the dependence of the Young and the shear modulus on the density of cross-links, note that ρ_p is the number of chains connecting

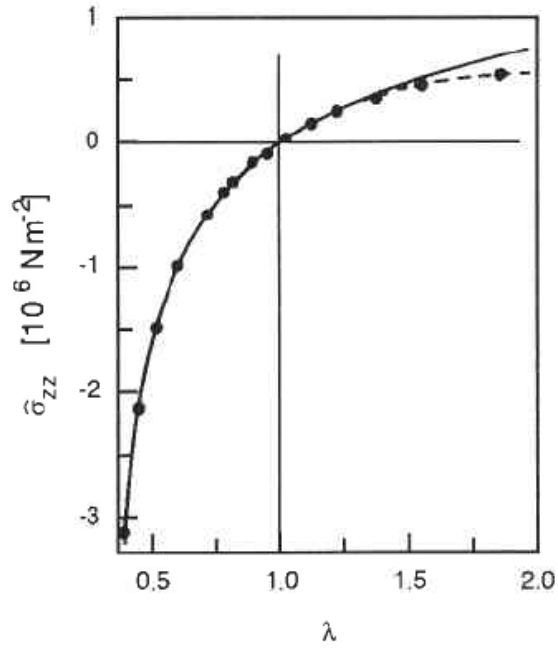


Figure 4.10: Stress vs. strain in rubber: Circles are experimental data for natural rubber and solid line is the theoretical prediction [Eq. (4.102)] [2].

two neighboring cross-links per unit volume. By increasing the density of cross-links ρ_p increases and thus E and μ increase too. Both E and μ can be expressed in terms of the average molecular mass between cross-links M_x . Note that $\rho_p = \tilde{\rho}/M_x$ where $\tilde{\rho}$ is the mass density of monomers and so

$$E = \frac{3\rho k_B T}{M_x} \quad (4.108)$$

and

$$\mu = \frac{\rho k_B T}{M_x}. \quad (4.109)$$

Chapter 5

Colloids

Colloids are heterogeneous systems consisting of particles of the dispersed phase distributed within the continuous phase. Colloids are ubiquitous in everyday life — many industrial, food, cosmetic, medical, and household products are colloidal dispersions. For example, mayonnaise is an emulsion of oil droplets, the primary emulsifier being egg yolk. Wall paint is a suspension of latex particles in water, and car paints are carefully engineered dispersions of pigments which provide the color, opacity, and a range of optical effects such as metallic finish. As the pigments are far more costly than the carrier continuous phase (typically a polymer solution which covers the surface so as to protect it from environmental effects), understanding how the colloidal dispersion is stabilized is of paramount importance.

Stability of colloidal dispersions is an essential problem in all colloidal systems. For example, putting oil and egg yolk in a bowl alone does not make mayonnaise. To disperse oil into small droplets, mechanical work is needed. This increases the surface energy of oil very much because the surface area of the dispersion is very big. Egg yolk and other additives are needed to stabilize the emulsion and to prevent the droplets from coalescing. — Let us estimate the total area of oil droplets in 100 g of mayonnaise. Assume that the radius of droplets is $\sim 1 \mu\text{m}$, that the density of oil is 1000 kg/m^3 , and that the droplets constitute all of the mayonnaise volume. (This cannot be true if all droplets are of the same size. The packing fraction of the closest packing of identical droplets is 74%.) Then the number of droplets is determined by $V = 4\pi NR^3/3$, where $V = 10^{-4} \text{ m}^3$ is the total volume. The area of droplets is

$$A = 4\pi NR^2 = \frac{3V}{R} = 300 \text{ m}^2. \quad (5.1)$$

As oil and egg yolk (about 50% water, 15% protein, and 35% fat) are immiscible, mechanical work is needed to mix them. This work is stored as the surface energy of the emulsion.

Generally speaking, both the continuous and the dispersed phase may be gas, liquid, or solid; the only exception to this scheme is the gas-gas combination because all gases mix. A dispersion of gas bubbles in a liquid is referred to as *foam* (e.g., whipped cream); in wet foam, the volume fraction of the continuous phase is considerable whereas in dry foam it is very small. Similarly, a dispersion of gas bubbles in a solid is called a *solid foam* (e.g., polyurethane foam). A dispersion of liquid droplets and small solid particles in a gas is an *aerosol* (e.g., mist) and a *solid aerosol* (e.g., smoke). A dispersion of liquid droplets in another immiscible liquid is an *emulsion* (e.g., mayonnaise) and a dispersion of solid particles in a liquid is a *sol* (e.g., blood). A dispersion of liquid particles in a solid matrix is a *gel* (e.g., butter and cheese) and solid particles in a solid matrix make a *solid sol* (e.g., colored glass, porcelain).

Key to the stability of colloidal dispersions is particle size, and this applies to many aspects of physics of colloids including sedimentation and hydrodynamics. The best way to demonstrate why size is so important is by comparing the order of magnitude of the typical energy scales [21]. Let us assume that the radius of the particles is $R = 1 \mu\text{m}$, their velocity is $v = 1 \mu\text{m/s}$, the viscosity of the continuous phase is $\eta = 10^{-3} \text{Ns/m}^2$, the density is $\rho = 1000 \text{kg/m}^3$, the relative density mismatch of the dispersed and the continuous phase is $\Delta\rho/\rho = 10^{-2}$, and the attraction energy of two particles in contact is $A_{eff} = 10^{-20} \text{J}$. These quantities are to be compared to each other and to the thermal energy $k_B T = 4 \times 10^{-21} \text{J}$ at room temperature.

We first compare the interparticle attraction to the thermal energy:

$$\frac{\text{attraction}}{\text{thermal energy}} = \frac{A_{eff}}{k_B T} \sim 1. \quad (5.2)$$

This tells us that colloidal dispersions are a thermal system described by classical (because the particles are much larger than atoms and molecules) statistical physics. Now compare the thermal energy to the work done by viscous forces at a distance equal to particle size:

$$\frac{\text{thermal energy}}{\text{viscous work}} = \frac{k_B T}{\eta v R^2} \sim 1. \quad (5.3)$$

This result suggests that viscous dissipation is very important. — Equally telling is the comparison of gravitational energy (the characteristic distance

is again particle size) and thermal energy:

$$\frac{\text{potential energy}}{\text{thermal energy}} = \frac{\Delta\rho R^4 g}{k_B T} \sim 10^{-1}. \quad (5.4)$$

Thus 1 μm size density-matched colloids do not sediment very much but in 10 μm particles sedimentation is very important. This shows that the line between the small, thermally agitated colloidal particles and the large, athermal granular systems is quite clear. Finally, we compare the inertial and the viscous forces:

$$\frac{\text{kinetic energy}}{\text{viscous work}} = \frac{\rho R^3 v^2}{\eta v R^2} = \text{Re} \sim 10^{-6}. \quad (5.5)$$

The Reynolds number is very small, indicating that the hydrodynamics of colloids is well within the Stokes flow regime.

The above estimates mean that colloids can be used as very convenient model systems to study a range of phenomena in statistical physics — and this can be done in fairly inexpensive tabletop experiments. The large size of particles is an important advantage because particles can be monitored and tracked using optical microscope.

5.1 Brownian motion*

Characteristic for colloids is Brownian motion. Incessantly hit by the molecules of the continuous phase, they move in a random fashion. This motion was first observed by Robert Brown who studied the jiggling trajectories of pollen in water. Since the collisions of water molecules with the particle are random the average force vanishes yet the particle diffuse from the starting point. The microscopic interpretation of the diffusion constant was proposed by Einstein in one of his 1905 papers; around the same time, a similar argument was worked out by Smoluchowski. Here we summarize Langevin's derivation of Einstein's result.

Consider the equation of motion of a spherical particle experiencing viscous damping described by the Stokes formula $\mathbf{F} = 6\pi\eta R\mathbf{v} = \xi\mathbf{v}$ and a random force \mathbf{F}_r due to collisions with the continuous phase:

$$m \frac{d^2\mathbf{r}}{dt^2} + \xi \frac{d\mathbf{r}}{dt} = \mathbf{F}_r. \quad (5.6)$$

As the x , y , and z directions are uncoupled and equivalent, we only need to consider one of them:

$$m \frac{d^2x}{dt^2} + \xi \frac{dx}{dt} = F_r, \quad (5.7)$$

where the F_r is the component of the random force along the x axis. We multiply both sides by x , rearrange, and replace

$$x \frac{d^2x}{dt^2} \quad \text{by} \quad \frac{d}{dt} \left(x \frac{dx}{dt} \right) - \left(\frac{dx}{dt} \right)^2. \quad (5.8)$$

Thus

$$\xi x \frac{dx}{dt} = xF_r - m \frac{d}{dt} \left(x \frac{dx}{dt} \right) + m \left(\frac{dx}{dt} \right)^2. \quad (5.9)$$

Note that the left-hand side can be written as

$$\xi x \frac{dx}{dt} = \frac{\xi}{2} \frac{dx^2}{dt}. \quad (5.10)$$

Now the presence of the two independent degrees of freedom in the equation of motion (position x and velocity v) is evident:

$$\frac{\xi}{2} \frac{dx^2}{dt} = xF_r - m \frac{d}{dt} (xv) + mv^2. \quad (5.11)$$

Upon averaging over the statistical ensemble, the products $\langle xF_r \rangle$ and $\langle xv \rangle$ vanish because x and F_r are uncorrelated and so are x and v . The only remaining nonzero terms left are

$$\frac{\xi}{2} \frac{d\langle x^2 \rangle}{dt} = m \langle v^2 \rangle. \quad (5.12)$$

According to the equipartition theorem, the right-hand side is equal to $k_B T$. Upon integrating both sides over t we have

$$\langle x^2 \rangle = \frac{2k_B T}{\xi} t \quad (5.13)$$

and so

$$\langle \mathbf{r}^2 \rangle = \frac{6k_B T}{\xi} t = 6Dt, \quad (5.14)$$

where we have defined the diffusion coefficient D . The importance of this result is that it relates D to the microscopic viscous force $\xi \mathbf{v}$ by the Stokes-Einstein relation:

$$D = \frac{k_B T}{6\pi\eta R}. \quad (5.15)$$

Thus one can measure the Boltzmann constant k_B by observing the Brownian motion of colloids. This was first done by Perrin in 1908 who used it to calculate the Avogadro number, thereby establishing atoms and molecules as real rather than virtual entities.

5.2 Interactions between colloids*

van der Waals attraction*

The most universal interaction between colloidal particles is the van der Waals interaction which arises from fluctuating dipolar moments of atoms. In a handwaving fashion, this interaction can be introduced as follows. Each atom is electrically neutral but at any given moment it does carry an instantaneous electric dipole because the centers of positive and negative charge coincide only on average. The magnitude of the instantaneous electric field of atom 1 \mathbf{E}_1 falls off proportional to r^{-3} , where r is the distance between the atoms. Due to the electric field of atom 1, atom 2 is polarized and the induced dipole moment is proportional to \mathbf{E}_1 at the location of atom 2: $\mathbf{p}_2 \propto \mathbf{E}_1(\mathbf{r}_2)$. The magnitude of \mathbf{p}_2 is proportional to r^{-3} . The energy of the polarized atom 2 in the field of atom 1 is

$$E_{int} = -\mathbf{p}_2 \cdot \mathbf{E}_1(\mathbf{r}_2) \propto r^{-6}, \quad (5.16)$$

the minus emphasizing that the interaction is attractive.

This simple argument exposes the main qualitative features of the van der Waals interaction. It also shows that this interaction is not pairwise additive: If there is a third atom close to atom 2, the latter is polarized both by atom 1 and atom 3. The induced dipole moment of atom 2 depends on the vectorial sum of electric fields of atoms 1 and 3, and thus the net interaction is not a simple algebraic sum of pair interactions. — An additional feature of the van der Waals interaction is that at large separations (e.g., beyond 10 nm or so) it decays $\propto r^{-7}$ due to retardation; in this regime, the finite speed of light is important.

The van der Waals interaction between atoms translates into interaction between colloidal particles. If the non-pairwise nature of atom-atom potential is neglected and if the atoms are distributed homogeneously, the resulting interaction between particles can be calculated fairly straightforwardly. For two halfspaces separated by a gap of width h , one finds [1]

$$E_{int} = -\frac{A_H A}{12\pi h^2}. \quad (5.17)$$

Here A is the area of the surfaces and A_H is the so-called Hamaker constant typically $\sim 10^{-19}$ J.

A complete analysis of the van der Waals interaction between colloidal particles covered by the Lifshitz theory is fairly complicated. Yet the mechanism leading to it can be appreciated by considering the so-called Casimir

effect, the interaction mediated between parallel metal plates by zero-point fluctuations of electromagnetic field.

5.2.1 Casimir interaction

The electromagnetic field between the plates can be represented by an ensemble of harmonic oscillators and at $T = 0$, these oscillators are all in ground state. The total energy reads

$$E(h) = 2 \sum_{\mathbf{k}} \frac{\hbar\omega(\mathbf{k})}{2}; \quad (5.18)$$

The prefactor of 2 accounts for the two polarizations. Due to boundary conditions at the plates, the transverse components of the field must vanish at $z = 0$ and at $z = h$, which restricts the allowed values of the transverse wavevector $k_z = n\pi/h$, where $n = 0, 1, 2, \dots$. The interaction energy is computed by recognizing that the energy Eq. (5.18) contains three terms: the bulk term is proportional to the volume between the plates, the surface term is proportional to the area of the plates, and the rest is the interaction. Recall that the force between the plates is defined as

$$\mathcal{F} = - \left(\frac{\partial E}{\partial h} \right)_{V,A}. \quad (5.19)$$

This definition is best visualized by thinking of the plates contained within a large container such that as the separation between them is changed the total volume of the container remains unchanged; in other words, as the plates are driven closer to each other the excess volume between them is "squeezed" into the space around the plates such that the total volume is conserved (Fig. 5.1).

To properly extract the interaction energy from the bare energy, the bulk term must be subtracted from Eq. (5.18); it turns out that in this particular case the surface term is zero. The bulk term is essentially the same as the bare energy except that the transverse wavevector is not discrete. The interaction energy reads

$$E_{int} = \hbar c \sum_{\mathbf{k}_{\perp}} \left[\sum'_{n=0}^{\infty} - \int_0^{\infty} dn \right] \sqrt{\mathbf{k}_{\perp}^2 + \left(\frac{n\pi}{h} \right)^2}; \quad (5.20)$$

the primed sum indicates that the $n = 0$ term is to be weighted by 1/2. Summation over the in-plane components of the wavevector can be replaced

Figure 5.1: Illustration of the definition of the structural force.

by an integral: $\sum_{\mathbf{k}_\perp} = [A/(2\pi)^2] \int_0^\infty 2\pi k_\perp dk_\perp = (A/2\pi) \int_0^\infty k_\perp dk_\perp$. Note that the bare energy and the bulk energy both diverge but as we will show below their difference does not.

The integral over k_\perp is calculated by introducing a damping function to make it finite. We first rewrite each term in Eq. (5.20) by introducing a new variable $u = \sqrt{\mathbf{k}_\perp^2 + (n\pi/h)^2}$:

$$E(n) = \frac{\hbar c A}{2\pi} \int_0^\infty k_\perp dk_\perp \sqrt{\mathbf{k}_\perp^2 + \left(\frac{n\pi}{h}\right)^2} = \frac{\hbar c A}{2\pi} \int_{n\pi/h}^\infty u^2 du. \quad (5.21)$$

To render the integral finite, we multiply the integrand by an exponential damping function:

$$E(n) = \frac{\hbar c A}{2\pi} \int_{n\pi/h}^\infty u^2 \exp(-u\delta) du \quad (5.22)$$

$$= \frac{\hbar c A \exp(-n\pi\delta/h)}{2\pi \delta^3} \left[\left(\frac{n\pi\delta}{h}\right)^2 + 2\frac{n\pi\delta}{h} + 2 \right]. \quad (5.23)$$

The difference between the sum and the integral on n in Eq. (5.20) can be calculated using the Euler-Maclaurin formula:

$$\sum_{n=1}^{m-1} E(n) = \int_0^m E(n) dn - \frac{1}{2} [E(0) + E(m)] + \frac{1}{12} \left. \frac{\partial E}{\partial n} \right|_0^m - \frac{1}{720} \left. \frac{\partial^3 E}{\partial n^3} \right|_0^m + \dots \quad (5.24)$$

Note that for finite $\delta > 0$ $E(\infty) = 0$. Thus we find that

$$E_{int} = \left[\sum_{n=0}^\infty ' - \int_0^\infty dn \right] E_n = \frac{1}{12} \left. \frac{\partial E}{\partial n} \right|_0^\infty - \frac{1}{720} \left. \frac{\partial^3 E}{\partial n^3} \right|_0^\infty. \quad (5.25)$$

The first derivative of $E(n)$ is $-(n^2\pi^3/h^3)\exp(-n\pi\delta/h)$ and vanishes both at $n = 0$ and at $n \rightarrow \infty$. The third derivative is proportional to $(-2 + 4n\pi\delta/h - n^2\pi^2\delta^2/h^2)\pi^3/h^3\exp(-n\pi\delta/h)$ which vanishes as $n \rightarrow \infty$ but at $n = 0$ it gives $2\pi^3/h^3$. Thus the interaction energy is

$$E_{int} = -\frac{\pi^2\hbar cA}{720h^3}. \quad (5.26)$$

The Casimir interaction is long-range and attractive; it is proportional to the area of the plates A and \hbar witnesses to its quantum-mechanical origin. (Note that taking the limit $\delta \rightarrow 0$ was not at all needed.)

We note in passing that at finite temperatures, thermal fluctuations become important. The form of the thermal Casimir interaction can be determined on dimensional grounds — it is natural to assume that it must be proportional to $k_B T$ rather than to $\hbar c$. Since $k_B T$ has the dimension of energy and not energy times length like $\hbar c$, we conclude that up to a multiplicative constant the thermal Casimir interaction must be given by

$$E_{int} \sim -\frac{k_B T A}{h^2}. \quad (5.27)$$

The main result of this scaling analysis is that the power law characteristic for the thermal Casimir force is different from that of the quantum Casimir force.

Electrostatic repulsion*

The van der Waals attraction is virtually omnipresent and although it decays rapidly with separation in most colloids it is quite strong at contact. On the other hand, most colloids are charged and the electrostatic repulsion between them is a means of stabilizing the particles from aggregating. Were the particles suspended in vacuum, the electrostatic repulsion would be described by the Coulomb law. But in a sol they are surrounded by a cloud of counterions which screen the particles' charge and makes their interaction more complicated. As the spatial distribution of the counterions depends on the position of the particles, the screened electrostatic repulsion is inherently non-pairwise additive.

Here we outline the simplest account of the screening known as the Poisson-Boltzmann theory. Consider a neutral solution of positive and negative ions of charge $+q$ and $-q$, respectively, in a position-dependent electric potential Ψ . In equilibrium, the density of each ion species depends on the

local potential and is determined by the canonical distribution

$$\rho_{\pm}(\mathbf{r}) = \rho_0 \exp\left(\mp \frac{q\Psi}{k_B T}\right), \quad (5.28)$$

where ρ_0 is the density of the ions in absence of particles. The potential itself must obey the Gauss law $\nabla \cdot \mathbf{E} = -\nabla^2 \Psi = \rho_e / (\epsilon \epsilon_0)$, where $\rho_e = q\rho_+ - q\rho_-$ is the charge density. In a simple 1D plate-plate geometry,

$$\epsilon \epsilon_0 \frac{d^2 \Psi}{dz^2} = q\rho_0 \left[\exp\left(\frac{q\Psi}{k_B T}\right) - \exp\left(-\frac{q\Psi}{k_B T}\right) \right] \quad (5.29)$$

or

$$\frac{d^2 \Psi}{dz^2} = \frac{2q\rho_0}{\epsilon \epsilon_0} \sinh\left(\frac{q\Psi}{k_B T}\right). \quad (5.30)$$

This is the Poisson-Boltzmann equation. If the potential is weak it can be linearized and

$$\frac{d^2 \Psi}{dz^2} = \frac{2q^2 \rho_0}{\epsilon \epsilon_0 k_B T} \Psi = \kappa^2 \Psi \quad (5.31)$$

is known as the Debye-Hückel approximation. Here

$$\kappa = \sqrt{\frac{2q^2 \rho_0}{\epsilon \epsilon_0 k_B T}} \quad (5.32)$$

is the inverse screening length.

Equation (5.31) can be easily integrated. In case of a single plate held at fixed potential, potential falls off exponentially:

$$\Psi(z) = \Psi_0 \exp(-\kappa z). \quad (5.33)$$

The exponentially decaying potential means that the density of counterions at the plate is increased relative to bulk whereas the density of coions is decreased (Fig. 5.2). This structure is often referred to as the *double layer*. As the ions have a finite size, the plate is effectively covered by a layer of counterions, on top of which is a (less compact) layer of coions (Fig. 5.3).

In case of weak potentials and large separations, the electrostatic interaction of two parallel plates of identical potential is short-range

$$E_{int} \propto \exp(-\kappa h). \quad (5.34)$$

Note that the range of the interaction depends on the concentration of ions. Adding salt to the solution increases the ion concentration, thereby decreasing the screening length κ^{-1} . This reduces the electrostatic repulsion between surfaces. — In the above illustration, we assumed that the potential on colloidal particles is fixed. Usually, the particles are characterized by a fixed charge rather than by fixed potential.

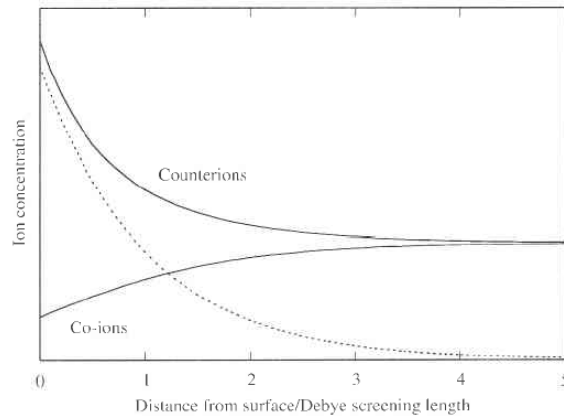


Figure 5.2: Screened electric potential at a plate held at fixed potential and the densities of counterions and coions [1].

Figure 5.3: Schematic of the electric double layer at a charged plate.

Depletion attraction*

Addition of polymers to a colloidal dispersion may generate either repulsive or attractive short-range forces. If the polymers adsorb on the colloids, the grafted polymer brushes keep the particles apart, preventing aggregation (Fig. 5.4). The range of the brush-brush repulsion is related to the degree of polymerization. The stabilization of colloids by adsorbed polymers happens in polymers in good solvents so that the chains themselves are not attracted to each other. In a poor solvent, polymer-polymer contacts are preferred and this leads to an attraction between particles.

If the polymers do not adsorb on the colloids, they give rise to a universal and chemistry-independent entropic interaction between the particles

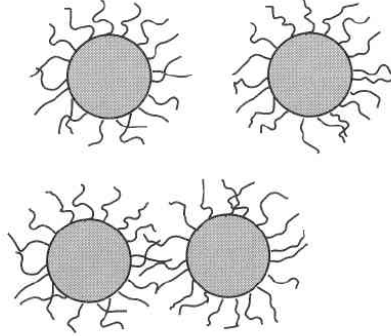


Figure 5.4: Polymers that adsorb on colloidal particles form brushes which prevent the colloids from aggregating [1].

known as the *depletion attraction*. First studied theoretically by Asakura and Oosawa in 1950s, the depletion attraction is caused by the overlap of the depletion zones around two nearby particles. In case of two parallel plates of area A , immersed in a suspension of polymer coils of diameter σ (treated as hard spheres for simplicity) and separated by more than σ , the excluded volume around each plate consists of a layer of thickness $\sigma/2$ and area A on either side of the plate. In the limit of very small density of polymer coils, the configuration integral reads

$$Z_N(h > \sigma) = (V - V_{excl})^N = (V - 2\sigma A)^N, \quad (5.35)$$

where we have included the total excluded volume of both depletion zones. At separations h smaller than σ the depletion zones between the plates overlap and the total excluded volume is decreased (Fig. 5.5). The configuration integral is

$$Z_N(h < \sigma) = [V - (\sigma + h)A]^N. \quad (5.36)$$

The separation-dependent part of the free energy of the coils is given by $F(h) = -k_B T \ln Z_N$ and like in the analysis of Casimir interaction, we define the interaction free energy relative to the reference state at large separations:

$$F_{depl} = F(h < \sigma) - F(h > \sigma) = -Nk_B T \ln \frac{V - (\sigma + h)A}{V - 2\sigma A}. \quad (5.37)$$

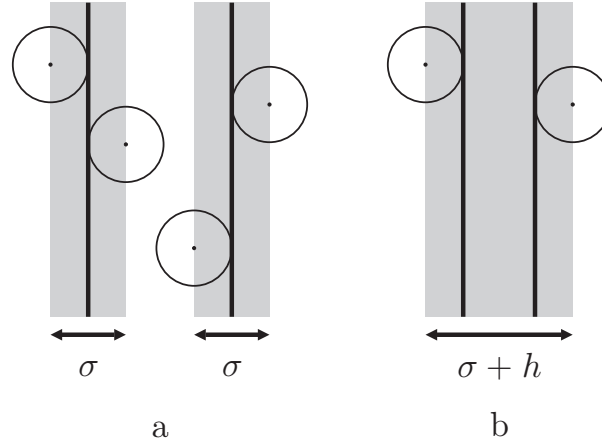


Figure 5.5: Depletion interaction between parallel plates. Shaded areas are the exclusion zones at large plate-plate separation $h > \sigma$ (a) and at small $h < \sigma$ (b).

As $V \gg A\sigma, Ah$, this can be approximated by

$$F_{depl} = -Nk_B T \ln \left(1 - \frac{(h - \sigma)A}{V - 2\sigma A} \right) \quad (5.38)$$

$$\approx -Nk_B T \frac{(\sigma - h)A}{V} = -\rho_p A k_B T (\sigma - h). \quad (5.39)$$

Here $\rho_p = N/V$ is the density of the polymer coils. — This result shows that the depletion interaction is attractive, proportional to the area of plates as expected, entropic, and to lowest order proportional to the density of polymers. The depletion force $\mathcal{F} = -\rho_p A k_B T$ is independent of separation for $h < \sigma$ and vanishes at $h > \sigma$, and can be interpreted as the osmotic pressure of the finite-size particles. At separations larger than σ , the polymer coils can enter the gap between the plates and the osmotic pressures on either side of each plate are balanced. On the other hand, for $h < \sigma$ the osmotic pressure within the gap is 0, and the outside osmotic pressure pushes the plates together.

In a similar fashion, one can calculate the interaction energy for the sphere-sphere geometry. Like in the case of plates, the range of interaction is given by σ and for $2R < h < 2R + \sigma$ (R being the radius of the large spheres) the result is

$$F_{depl} = -\frac{\pi}{4} \rho_p k_B T (2R + \sigma - h)^2 (2R + \sigma + h/2). \quad (5.40)$$

In a more complete theory of depletion interaction, one must account for the finite density of the polymers. At finite density, the structure of the fluid of small particles within the gap is important. As the density is increased the particles become increasingly more positionally ordered and the depletion interaction develops oscillations.

We also stress that the depletion interaction does not act only between the colloids but also between the colloids and a substrate. It is not hard to see that the overlap of depletion zones of a spherical particle and a flat substrate is larger than that of two spherical particles (Fig. 5.6). Moreover, on a grooved substrate the attraction is largest within the grooves and smallest on the ridges.

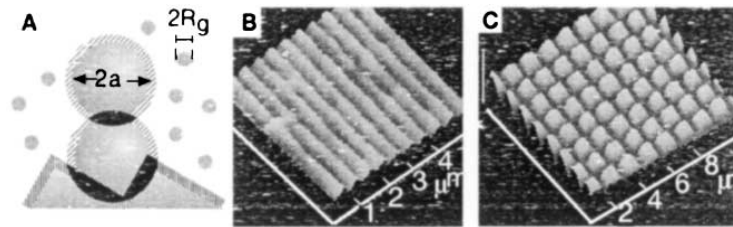


Figure 5.6: Depletion attraction also gives rise to attraction of colloids to substrates. On a grooved substrate, particles are driven to corners and they avoid ridges (a) [23]. A patterned substrate can serve as a template (b, c).

Derjaguin approximation

In many cases, the interaction energy is most easily evaluated in the plate-plate geometry but experimentally the sphere-sphere and crossed-cylinder geometries are far more interesting: The former because it is characteristic of typical colloids and the latter because it is much more convenient for direct force measurements than the plate-plate geometry where it is difficult to ensure that the plates are really parallel. The interaction potential in the plate-plate geometry can be related to the force between two spheres using the Derjaguin approximation valid for any force law provided that the range of the interaction is short compared to sphere diameter, and it can be used for van der Waals, electrostatic, and depletion interactions.

The Derjaguin approximation states that

$$\mathcal{F}(h) \approx 2\pi \frac{R_1 R_2}{R_1 + R_2} \frac{E_{int}(h)}{A}, \quad (5.41)$$

where $\mathcal{F}(h)$ is the force between spheres of radii R_1 and R_2 separated by a gap of width h and $E_{int}(h)$ is the interaction energy of two plates h apart.

5.3 Derjaguin-Landau-Verwey-Overbeek theory*

These interactions are essential for the understanding of colloidal stability. The van der Waals interaction alone would cause the particles to stick together. The structure of thus formed aggregates is determined by the kinetics of their formation. The aggregates themselves are not in thermodynamic equilibrium and typically consist of branching fractal clusters (Fig. 5.7). In the diffusion-limited aggregation, particles stick to each other upon contact so that as soon as a new particle touches an existing cluster it becomes immobile. This prevents them from reaching the minimal-energy configuration. If the sticking rate is not determined by diffusion, one speaks of reaction-limited aggregation. In this regime particles have ample opportunities to explore a given local neighborhood.

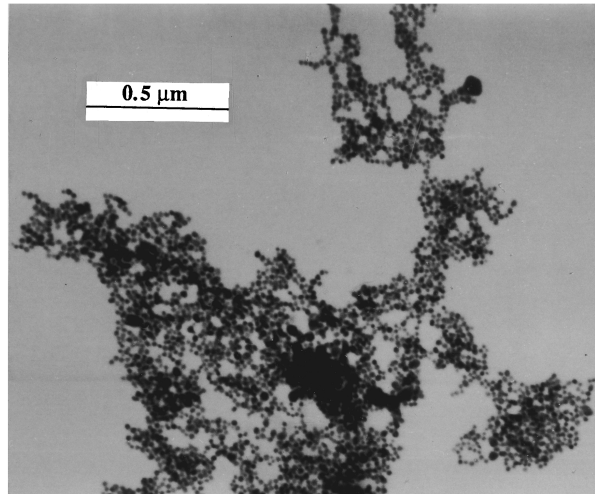


Figure 5.7: Disordered colloidal aggregate typical for diffusion-limited aggregation. Electron micrograph of silver colloidal particles [24].

To prevent the formation of the non-equilibrium aggregates, colloids must be stabilized. This can be done either by covering the particles by a polymer brush or by electrostatic stabilization. The classical theory of colloidal stabilization is due to Derjaguin, Landau, Verwey, and Overbeek (DLVO). In this theory, the total pair potential between particles consists

of the van der Waals attraction and the electrostatic repulsion. The absolute minimum of the interaction is at close contact (Fig. 5.8) but if the electrostatic potential is not screened the energy barrier due to repulsion is very high compared to $k_B T$. As salt is added and the screening length is decreased, the electrostatic repulsion is less and less prominent and a secondary minimum develops at a finite separation. If the depth of this minimum is $\sim k_B T$, particles behave like a thermal ensemble, exploring all of the phase space. If the screening length is further decreased, the barrier separating the absolute minimum at contact vanishes and the colloids enter the aggregation regime.

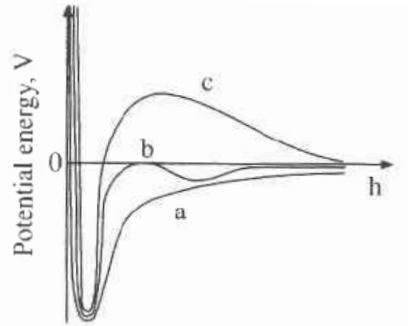


Figure 5.8: DVLO theory of colloidal stability: Electrolyte concentration is largest in curve a and smallest in curve c [9].

5.4 Phase diagram of hard-sphere colloids

Stabilized colloids readily form a range of ordered and disordered phases. The paradigmatic hard-sphere colloidal system is fluid at small packing fractions and forms an entropic FCC crystal (Fig. 5.9). The phase diagram depends only on packing fraction:

$$\text{fluid} \begin{array}{c} \longleftrightarrow \\ 0.494 \end{array} \text{coexistence} \begin{array}{c} \longleftrightarrow \\ 0.545 \end{array} \text{FCC crystal}$$

Theoretically, FCC crystal is stable up to a packing fraction of 0.740 where spheres reach close packing. In experiments and in simulations a glass (= structurally disordered arrested) phase is found very often at packing fractions between 0.58 and 0.644; packing fraction of 0.644 corresponds to the so-called random close-packed structure.

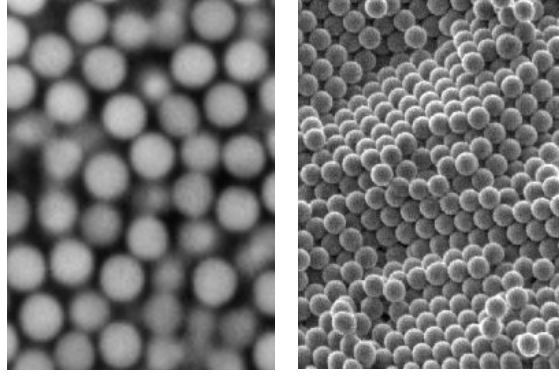


Figure 5.9: Colloidal hard-sphere fluid (left) and FCC crystal (right; A. van Blaaderen).

The fluid-FCC phase transition and the stability of the FCC crystal can be interpreted in terms of excluded volume. The free energy of the fluid phase can be estimated using the Carnahan-Starling equation of state [Eq. (2.93)]. Note that

$$p = - \left(\frac{\partial F}{\partial V} \right)_{T,N} = \frac{\rho^2}{N} \left(\frac{\partial F}{\partial \rho} \right)_{T,N} \quad (5.42)$$

which can be integrated to calculate the excess free energy

$$F_l^{ex} = Nk_B T \int_0^\rho \frac{d\rho}{\rho} \frac{\beta p^{ex}}{\rho} = Nk_B T \int_0^\eta \frac{d\eta}{\eta} \left[\frac{1 + \eta + \eta^2 - \eta^3}{(1 - \eta)^3} - 1 \right] \quad (5.43)$$

$$= Nk_B T \frac{4\eta - 3\eta^2}{(1 - \eta)^2}. \quad (5.44)$$

The excess chemical potential is given by $\beta\mu_l^{ex} = (F^{ex} + pV)/(Nk_B T) = F^{ex}/(Nk_B T) + \beta p^{ex}/\rho$ or

$$\beta\mu_l^{ex} = \frac{1 + 5\eta - 6\eta^2 + 2\eta^3}{(1 - \eta)^3} \quad (5.45)$$

and by adding it to the total chemical potential is

$$\beta\mu_l = \ln \eta - 1 + \beta\mu_l^{ex} = \ln \eta + \frac{8\eta - 9\eta^2 + 3\eta^3}{(1 - \eta)^3} \quad (5.46)$$

[On inserting Eq. (2.16), we have dropped the additive constant $3 \ln \Lambda$.]

Now we estimate the free energy of the FCC crystal. In the cell model, we assume that each particle is confined to the cage formed by its neighbors, and its free volume is $\sim \left(V^{1/3} - V_{cp}^{1/3}\right)^3$. Here the numerical prefactor depends on the shape of the cell (i.e., on the symmetry of the lattice), V is the volume of the unit cell containing 1 particle so that the packing fraction $\eta = \pi\sigma^3/(6V)$, and V_{cp} is the volume of the unit cell at close packing. [An illustration: In 2D square lattice, the unit cell is a square of size $a - \sigma$ and the free 2D volume is $(a - \sigma)^2$ (Fig. 5.10).] The free energy then reads

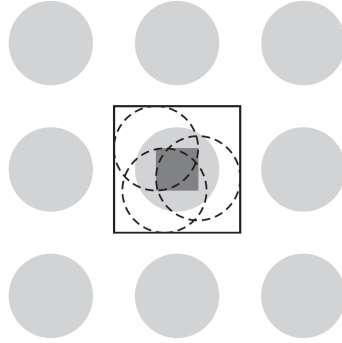


Figure 5.10: 2D free volume in square lattice: Solid line is the "cage" formed by the neighbors, and the center of mass of the reference particle may explore the dark area in its center. This is the 2D free volume.

$$F_s = -Nk_B T \ln V_{free} = -3Nk_B T \ln \left(V^{1/3} - V_{cp}^{1/3} \right) + const. \quad (5.47)$$

where the additive constant depends on the symmetry of the lattice in question. The pressure is

$$p_s = \frac{Nk_B T}{V \left[1 - (V_{cp}/V)^{1/3} \right]} \quad (5.48)$$

so that

$$\frac{\beta p_s}{\rho} = \frac{1}{1 - (\eta/\eta_{cp})^{1/3}} \approx \frac{3}{1 - \eta/\eta_{cp}}, \quad (5.49)$$

where $\eta = \pi\sigma^3/V$ and $\eta_{cp} = \pi\sigma^3/V_{cp}$. In the last step, we have used an approximation suitable for large packing fractions $\eta \rightarrow \eta_{cp}$ which gives upon integration

$$F_s \approx Nk_B T \left[3 \ln \frac{\eta}{1 - \eta/\eta_{cp}} + 2.1178 \dots \right]; \quad (5.50)$$

the numerical constant corresponds to the FCC lattice where $\eta_{cp} = 0.74$.

Now we can calculate the chemical potential

$$\beta\mu_s \approx 3 \ln \frac{\eta}{1 - \eta/\eta_{cp}} + \frac{3}{1 - \eta/\eta_{cp}} + 2.1178 \dots \quad (5.51)$$

To determine the location of the phase transition, we plot the chemical potentials of the hard-sphere liquid and FCC solid [Eqs. (5.46) and (5.51)] against their respective pressures [Eqs. (2.93) and (5.49)]; these equations need to be multiplied by $\eta = \pi\sigma^3\rho/6$ so that their left-hand sides do not include packing fraction. The result is shown in Fig. 5.11 and at the transition, both the chemical potentials and the pressures must coincide. This happens at $\eta_l = 0.491$ and $\eta_s = 0.541$ — very close to the predictions of more sophisticated theories.

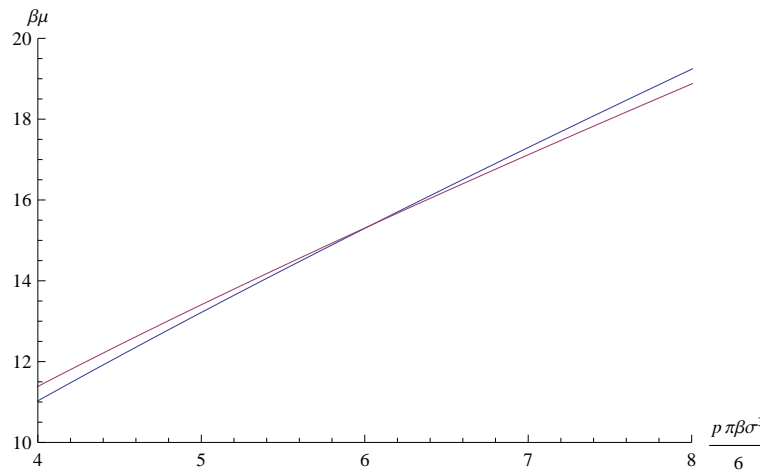


Figure 5.11: Chemical potentials of hard-sphere liquid (blue) and FCC crystal (magenta). Beyond $\pi\sigma^3\beta p/6 \approx 6.08$, the FCC crystal is stable.

This analysis of the phase diagram of hard spheres is just the simplest example of the phase behavior of colloids. Hard colloidal particles of anisotropic shape (rods and platelets) readily form liquid-crystalline phases; some of this behavior was covered in Ch. 3. In soft colloids, the condensed part of the phase diagram may contain more complex lattices than FCC.

Chapter 6

Amphiphiles

Amphiphile molecules consist of a polar hydrophilic part, typically a head, and an apolar hydrophobic part, typically an alkyl chain. When dispersed in water, their dual nature causes the amphiphiles to self-organize such that the polar parts face water (thereby reducing the surface tension of water) whereas the apolar parts orient away from water. Many substances exhibit amphiphilic behavior — the membranes of animal cells and cell organelles are based on phospholipids, fatty acids consisting of a polar head, usually a phosphoglyceride, and two hydrophobic alkyl chains of 10 to 20 C atoms. Amphiphilic behavior is also seen in diblock copolymers, polymers consisting of incompatible blocks joined by a covalent bond and dispersed in a solvent compatible with one block but not with the other.

When dispersed in water, amphiphiles self-assemble in various ways depending on the shape of the molecules. The key geometrical parameter is the ratio

$$\frac{v}{l_c a_0}, \quad (6.1)$$

where v is the volume of the molecules, l_c is the length of the apolar chain, and a_0 is the area per headgroup. If the amphiphile has a single short tail and if the area of the head is large, then the amphiphile looks like an ice-cream cone. If

$$\frac{v}{l_c a_0} \leq \frac{1}{3}, \quad (6.2)$$

amphiphiles will pack into spherical micelles (Fig. 6.1). In truncated-cone amphiphiles

$$\frac{1}{3} \leq \frac{v}{l_c a_0} \leq \frac{1}{2}, \quad (6.3)$$

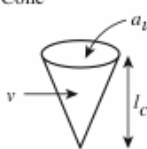


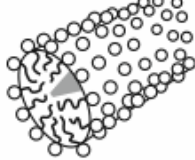


Lipid	Critical packing parameter v/a_0l_c	Critical packing shape	Structures formed
Single-chained lipids (surfactants) with large head-group areas: <i>SDS in low salt</i>	$<1/3$	Cone 	spherical micelles 
Single-chained lipids with small head-group areas: <i>SDS and CTAB in high salt, non-ionic lipids</i>	$1/3-1/2$	Truncated cone 	cylindrical micelles 
Double-chained lipids with large head-group areas, fluid chains: <i>phosphatidyl choline (lecithin), phosphatidyl serine, phosphatidyl glycerol, phosphatidyl inositol, phosphatidic acid, sphingomyelin, DGDG^a, dihexadecyl phosphate, diakyl dimethyl ammonium salts</i>	$1/2-1$	truncated cone 	flexible bilayers, vesicles 

Figure 6.1: Morphologies of amphiphile self-assemblies in water [20]. For $v/(l_c a_0) > 1$ amphiphiles form inverted cylindrical and micellar structures. Also possible are more complex bicontinuous structures.

and they form cylindrical micelles. For

$$\frac{1}{2} \leq \frac{v}{l_c a_0} \leq 1 \quad (6.4)$$

the amphiphiles self-organize in bilayers and lamellar phases, whereas for

$$\frac{v}{l_c a_0} > 1 \quad (6.5)$$

they form inverted micelles, which can be either cylindrical or spherical. This is how the shape of the phospholipids, which is controlled by their chemical structure, generates the curvature of phospholipid self-assemblies.

6.1 Amphiphile aggregation*

In water suspension, amphiphiles will form a given type of aggregate only if the concentration is large enough. In equilibrium, the chemical potential of dissolved single molecules must be equal to the chemical potential of molecules in an aggregate. The critical concentration depends on the shape of amphiphiles, and so does the size distribution of the aggregates.

Let us derive the equilibrium mole fraction of amphiphiles in aggregates containing N molecules. The total number of amphiphiles in N -aggregates is N_N , and the number of N -aggregates is $n_N = N_N/N$. The chemical potential of N -aggregates is

$$\tilde{\mu}_N = \tilde{\mu}_N^{0,\rho} + k_B T \ln \rho_N, \quad (6.6)$$

where $\rho_N = n_N/V$ is the (number) density of N -aggregates. This expression can be rewritten in terms of the mole fraction of the N -aggregates:

$$X_N = \frac{N_N}{N_{tot}}, \quad (6.7)$$

where $N_{tot} = N_w + N$ is the total number of water and amphiphile molecules; if the concentration of amphiphiles is small, $N_{tot} \approx N_w$. Note that $\sum_N X_N$ equals the total mole fraction of amphiphiles $X \neq 1$. — Hence

$$X_N = \frac{N n_N}{N_{tot}} = \frac{N n_N/V}{N_{tot}/V} \approx \frac{N \rho_N}{\rho_w} \quad (6.8)$$

so that $\rho_N \approx \rho_w X_N/N$

$$\tilde{\mu}_N = \tilde{\mu}_N^0 + k_B T \ln \left(\frac{X_N}{N} \right), \quad (6.9)$$

where $\tilde{\mu}_N^0$ is the reference chemical potential of the whole N -aggregate. The chemical potential per amphiphile in the N -aggregate is

$$\mu_N = \frac{\tilde{\mu}}{N} = \mu_N^0 + \frac{k_B T}{N} \ln \left(\frac{X_N}{N} \right). \quad (6.10)$$

In equilibrium, μ_N must be equal to μ_1 , the chemical potential of dissolved amphiphiles:

$$\mu_N^0 + \frac{k_B T}{N} \ln \left(\frac{X_N}{N} \right) = \mu_1^0 + k_B T \ln (X_1). \quad (6.11)$$

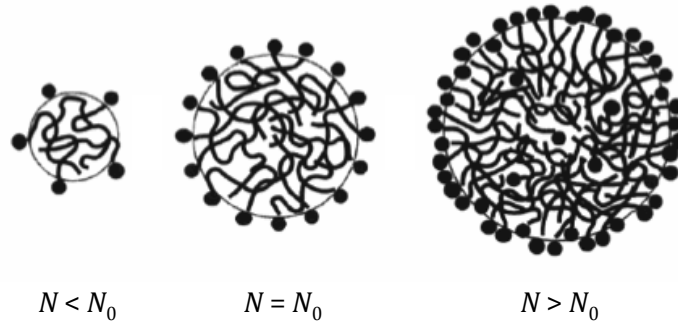


Figure 6.2: Spherical micelles: In micelles that are either too small ($N < N_0$) or too big ($N > N_0$), amphiphile chemical potential is larger than in optimal micelles [1].

Upon rearranging we obtain the law-of-mass action

$$\frac{X_N/N}{X_1^N} = \exp\left(\frac{N(\mu_1^0 - \mu_N^0)}{k_B T}\right) \quad (6.12)$$

and

$$X_N = N \left[X_1 \exp\left(\frac{\mu_1^0 - \mu_N^0}{k_B T}\right) \right]^N. \quad (6.13)$$

This important result shows that if $\mu_1^0 < \mu_N^0$, most amphiphiles are dissolved and $X_{N>1}$ will be small whereas if $\mu_1^0 > \mu_N^0$ aggregates are readily formed. Equation (6.13) will now be used to derive the distribution of aggregate size for spherical micellar, cylindrical micellar, and bilayer aggregates.

Spherical micelles*

At a given headgroup area a_0 and tail length l_c , the optimal radius of a spherical micelle is l_c and the optimal micelle area is $4\pi l_c^2$ so that the optimal number of amphiphiles is $N_0 = 4\pi l_c^2/a_0^2$ (Fig. 6.2). In micelles that are too small, the area per headgroup is too big and in micelles that are too big some of the amphiphiles invariably end up buried in the hydrophobic core. Thus we can assume that the chemical potential has a minimum at N_0 :

$$\mu_N^0 = \mu_{N_0}^0 + \Lambda(N - N_0)^2. \quad (6.14)$$

Upon inserting this result into Eq. (6.13) and a slight rearrangement, we find that the distribution of micelle size is almost Gaussian:

$$X_N = N \left[\frac{X_{N_0}}{N_0} \exp \left(\frac{-N_0 \Lambda (N - N_0)^2}{k_B T} \right) \right]^{N/N_0}. \quad (6.15)$$

In case that the minimum of Eq. (6.14) is deep compared to $k_B T$, the only relevant states of the amphiphiles are the optimal micelles and dissolved amphiphiles. Let us compare their respective concentrations X_{N_0} and X_1 :

$$X_{N_0} = N_0 \left[X_1 \exp \left(\frac{\mu_1^0 - \mu_{N_0}^0}{k_B T} \right) \right]^{N_0}. \quad (6.16)$$

If the concentration of dissolved amphiphiles is smaller than the *critical micelle concentration*

$$X_c = \exp \left(-\frac{\mu_1^0 - \mu_{N_0}^0}{k_B T} \right), \quad (6.17)$$

then virtually all amphiphiles are dissolved and the concentration of micelles vanishes. [Note that the total concentration of amphiphiles is $X_1 + X_{N_0}$.] On the other hand, for $X_1 > X_c$ the concentration of amphiphiles self-organized in the micelles grows fairly rapidly whereas the concentration of dissolved amphiphiles levels off at a constant value.

Cylindrical micelles

Amphiphiles with $v/(l_c a_0)$ between 1/3 and 1/2 form cylindrical micelles which are qualitatively different from the spherical micelles in that there is no preferred size. Most amphiphiles in cylindrical pack consistently with their truncated-cone shape, the only exception being the two hemispherical caps which cost an extra energy ΔE conveniently expressed in terms of the thermal energy $\alpha k_B T$, α being a numerical constant. The average chemical potential per molecule reads

$$\mu_N^0 = \mu_\infty + \frac{\alpha k_B T}{N}, \quad (6.18)$$

where μ_∞ is the chemical potential in an infinite micelle. For simplicity, we extend the validity of this formula down to $N = 1$ corresponding to dissolved micelles. By inserting it in Eq. (6.13) we find that

$$X_N = N [X_1 \exp(\alpha(1 - 1/N))]^N = N [X_1 \exp \alpha]^N \exp(-\alpha). \quad (6.19)$$

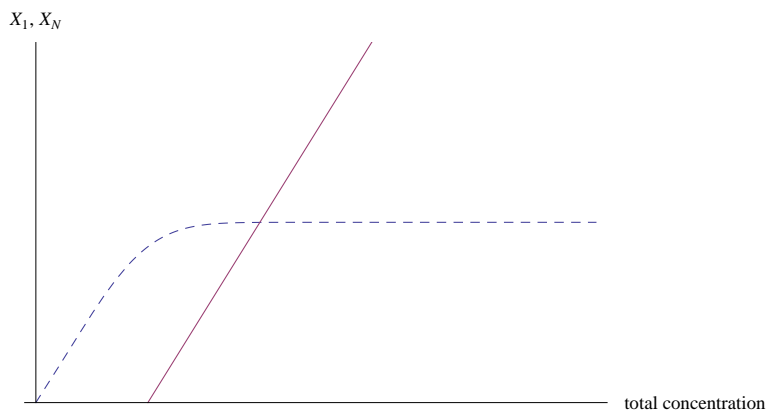


Figure 6.3: Dependence of the concentration of spherical micelles (solid line) and dissolved molecules (dashed line) on the overall concentration of amphiphiles. Below critical micelle concentration, micelles are absent whereas beyond it the concentration of dissolved molecules saturates.

X_N is finite for $N \rightarrow \infty$ only if X_1 reaches $\sim \exp(-\alpha)$; this corresponds to the critical micelle concentration.

Now we can derive the distribution of micelle size. First we calculate the sum of concentrations of all micelles. Note that $\sum_{N=1}^{\infty} N y^N = \sum_{N=0}^{\infty} N y^N = y(d/dy) \sum_{N=0}^{\infty} y^N = y/(1-y)^2$. Thus

$$X_{micelles} = \sum_{N=2}^{\infty} X_N = \sum_{N=1}^{\infty} X_N - X_1 = \frac{X_1}{(1 - X_1 \exp \alpha)^2} - X_1. \quad (6.20)$$

We have learned from the analysis of the formation of spherical micelles that beyond the critical micelle concentration where $X_{micelles} \exp \alpha \gg 1$, X_1 saturates at $\exp(-\alpha)$. So we write $X_1 = \exp(-\alpha) - b$ and keep the lowest-order term in b to find that well beyond the critical micelle concentration $b = 1/\sqrt{X_{micelles} \exp 3\alpha}$ and

$$X_1 \approx \left(1 - \frac{1}{\sqrt{X_{micelles} \exp \alpha}}\right) \exp(-\alpha). \quad (6.21)$$

Hence

$$X_N = N \left(1 - \frac{1}{\sqrt{X_{micelles} \exp \alpha}}\right)^N \exp(-\alpha). \quad (6.22)$$

Unlike in spherical micelles, the size distribution of cylindrical micelles is very broad, characterized by a long tail (Fig. 6.4). The distribution peaks at $N_{max} = \sqrt{X_{micelles} \exp \alpha}$ [1].

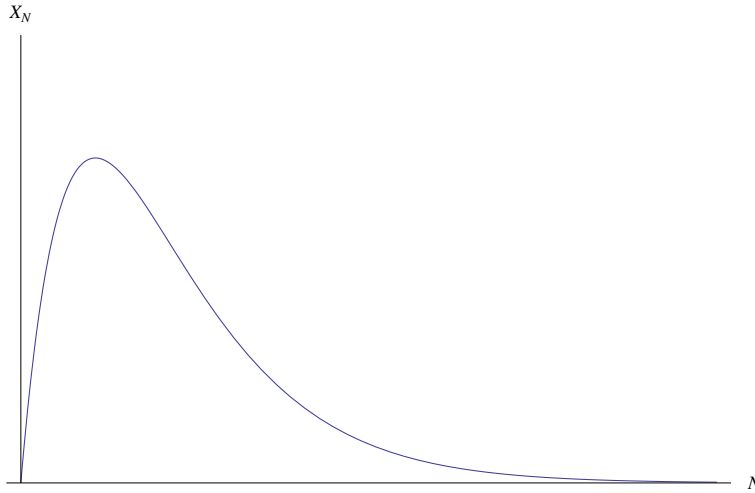


Figure 6.4: Size distribution of cylindrical micelles.

Bilayer membranes

Amphiphiles of cylindrical shape self-organize in bilayers and their aggregation behavior is qualitatively different from that of micelle-forming amphiphiles. Consider a finite bilayer patch — a disk of N molecules of area $\sim N$ and radius $\sim \sqrt{N}$. In such a patch, all amphiphiles except those in the rim are packed optimally consistent with their cylindrical shape. The number of sub optimally packed molecules is proportional to the length of the approximately hemicylindrical rim, i.e., $\sim \sqrt{N}$. Thus the average chemical potential per molecule is

$$\mu_N^0 = \mu_\infty + \frac{\alpha k_B T}{\sqrt{N}}, \quad (6.23)$$

which means that

$$X_N = N \left[X_1 \exp \left(\alpha \left(1 - 1/\sqrt{N} \right) \right) \right]^N = N [X_1 \exp \alpha]^N \exp(-\alpha N^{1/2}). \quad (6.24)$$

At large N , the number of finite disklike patches decreases exponentially because of the factor $\exp(-\alpha N^{1/2})$, which implies that the dissolved amphiphiles must be in equilibrium with an infinite bilayer membrane.

This can be seen as follows. Assume that a finite planar membrane patch is in equilibrium with isolated dissolved amphiphiles as well as with

an infinite membrane. In this case,

$$\mu_\infty = \mu_1 = \mu_N \quad (6.25)$$

so that

$$\mu_\infty = \mu_1^0 + k_B T \ln X_1 \quad (6.26)$$

and

$$\mu_\infty = \mu_N^0 + \frac{k_B T}{N} \ln \frac{X_1}{N}. \quad (6.27)$$

These two equations can be inverted to give

$$X_1 = \exp\left(-\frac{\mu_1^0 - \mu_\infty}{k_B T}\right) \quad (6.28)$$

and

$$X_N = N \exp\left(-\frac{N(\mu_N^0 - \mu_\infty)}{k_B T}\right) = N \exp(-\alpha N^{1/2}). \quad (6.29)$$

In the last step we used Eq. (6.23).

Now the total concentration of amphiphiles in all finite patches including isolated dissolved molecules is

$$X = X_1 + \sum_{N=2}^{\infty} X_N = \exp\left(-\frac{\mu_1^0 - \mu_\infty}{k_B T}\right) + \sum_{N=2}^{\infty} N \exp(-\alpha N^{1/2}) \quad (6.30)$$

Note that the sum converged to a finite value so that X itself is finite. Thus as more amphiphiles are added to the solution, they must be included in the infinite membrane.

In case of cylindrical micelles, the situation is qualitatively different. Here

$$X = X_1 + \sum_{N=2}^{\infty} N \exp(-\alpha) \quad (6.31)$$

which does not converge. We conclude that finite cylindrical micelles cannot coexist with infinite cylindrical micelle.

A large membrane may be closed so as to eliminate the energy cost of the rim. Closed membranes are referred to as vesicles, and vesicle shape is determined by membrane elasticity.

6.2 Membrane elasticity

Much like in the case of surface energy of simple liquids, the microscopic intermolecular interactions between amphiphiles organized in a bilayer membrane give rise to the membrane stretching and bending energy. Here we derive this energy for a symmetric membrane consisting of monolayer leaflets each containing N molecules. The central quantity of interest is the average free energy per molecule, which depends on the deviation of the membrane area from the equilibrium area and on the deviation of the membrane shape from the flat undeformed shape.

The arguments presented below are based on three assumptions:

- **Monolayers are 2D fluids** This implies that the molecules can rearrange within the monolayers so as to minimize their mutual interaction energy, which means that the energy depends only on the density and not on the exact in-plane positional order like in crystalline membranes. Thus the free energy per molecule can be expressed in terms of the average membrane area per amphiphile headgroup a .
- **Membrane thickness is constant** The length of tails does not change much as the membrane is stretched or bent, and thus the distance between the two amphiphile-water interfaces remains constant. Specifically, $h = 2l$ where l is the length of amphiphiles.
- **Harmonic expansion** For small deformations, the area per headgroup can be expanded in terms of relative membrane area change and membrane curvature.

The total free energy consists of i) surface energy associated with the headgroup-water interface, ii) headgroup free energy associated with the interaction among headgroups, and iii) the free energy of tails:

$$F = F_s + F_h + F_t. \quad (6.32)$$

Per molecule, the free energy reads

$$f(a, h) = \frac{F(N, a, h)}{2N}. \quad (6.33)$$

Each of the three terms in $f(a, h)$ is characterized by a separate dependence on area per headgroup a .

The bilayer exists because the apolar tails of the amphiphiles do not mix with water. Thus we may expect that there is an energy cost for increasing the area of the amphiphile-water interface. The surface energy per headgroup is described by

$$f_s = \gamma a. \quad (6.34)$$

Typical values of the surface tension are $\gamma \approx 0.05 \text{ J/m}^2$ or about $0.1k_B T/\text{nm}^2$.

The interaction among headgroups is more complicated and includes electrostatic repulsion due to the polar nature of the headgroups and steric repulsion between them. Both of these terms can be modeled by [26]

$$f_h = \frac{W}{a}, \quad (6.35)$$

where W is a constant.

Apart from the head-head repulsion, amphiphiles repel each other because of the interaction of their tails. When pushed against each other, their tails must be stretched such or else the volume per amphiphile $v = al$ would not be conserved. This increases the energy of the tails and if we assume that the tails can be described by freely-jointed polymer chains of equilibrium length $l_c \propto n^{1/2}$ and volume $v = al \propto n$, we can write

$$f_t \propto \left(\frac{l}{l_c}\right)^2 \propto \frac{v^2}{a^2 n} \propto \frac{n}{a^2}. \quad (6.36)$$

A detailed analysis of the tail-tail repulsion and the ensuing lateral pressure within the membrane is quite complicated [26] and inessential for this discussion.

Irrespective of the exact form of the cohesive and the repulsive terms in the free energy, its dependence on the area per headgroup around equilibrium is harmonic (Fig. 6.5):

$$f(a, h) = f_0 + \frac{1}{2} \frac{d^2 f}{da^2} (a - a_0)^2 \quad (6.37)$$

Area compressibility

We first calculate the energy penalty for membrane stretching or compression. In this case, the area of headgroup of all molecules changes by the same amount and the variation of free energy is given by

$$\delta f = \frac{1}{2} \frac{\partial^2 f}{\partial a^2} (a - a_0)^2 = \frac{1}{2} \kappa_{AA} a_0 \left(\frac{\delta a}{a_0}\right)^2. \quad (6.38)$$

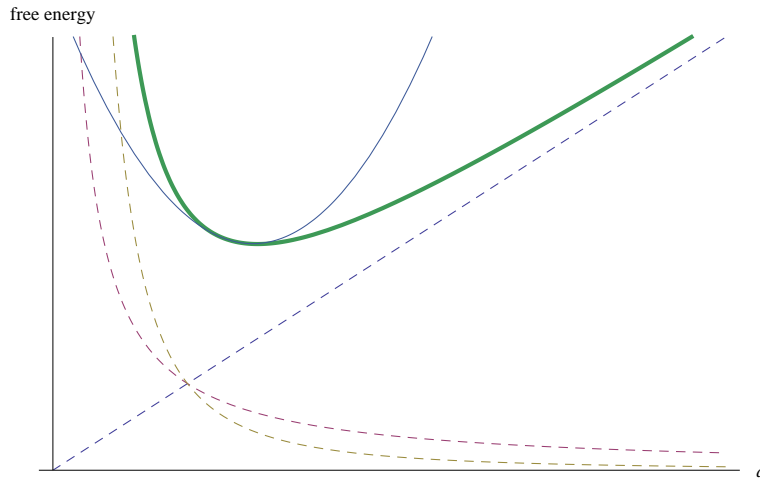


Figure 6.5: Free energy of amphiphiles in a bilayer: Surface energy (blue), head-head repulsion (magenta), and tail-tail repulsion (tan). Around equilibrium at $a = a_0$, the total free energy (thick green line) is a harmonic function of $a - a_0$.

The area compressibility modulus κ_A is defined as $a_0(\partial^2 f / \partial a^2)_{a_0}$. As the surface free energy is linear in a it does not contribute to κ_A which turns out to be dominated by the tail term [26]; typical values are $\kappa_A \sim 50k_B T/\text{nm}^2$.

Bending elasticity

Let us now see how the free energy of a bilayer changes upon bending. In this case, the variations of the area per headgroup of molecules in monolayers A and B are not the same. To derive the functional form of the bending energy, we need to know how does the area of the monolayers depend on the shape of the membrane; more precisely, how does it depend on its curvature.

The easiest way to introduce the curvature of a surface is by relying on the curvature of planar curves

$$C = \mathbf{n} \cdot \frac{\partial \mathbf{t}}{\partial s} = \mathbf{n} \cdot \frac{\partial^2 \mathbf{r}}{\partial s^2}. \quad (6.39)$$

Here $\mathbf{r}(s)$ is the location of a point on a curve parametrized by the arclength s , and $\mathbf{t}(s)$ is the tangent. On a surface, we can move along many paths which are planar curves. For each of these paths, the curvature can be defined by Eq. (6.39). Obviously, the curvature of any path is specific for

that path but it also reflects the shape of the surface. It turns out that for any surface there exist two perpendicular paths such that the curvatures along these paths are extremal; one is the smallest of all curvatures and the other one is the largest. These two curvatures are called the principal curvatures and denoted by C_1 and C_2 . In a cylinder of radius R , $C_1 = 0$ and $C_2 = 1/R$.

Curvature can be either positive or negative. Conventionally, we use the outward normal and so the curvature of a filled cylinder is positive and the curvature of a cylindrical hole is negative (in both cases, we talk about the curvature in the plane normal to cylinder axis). In a sphere, $C_1 = C_2 = 1/R$. For a saddle-shaped surface, $C_1 = 1/R_1$ and $C_2 = -1/R_2$.

Instead of describing the local curvature with the principal curvatures, we can use the mean curvature

$$H = \frac{1}{2}(C_1 + C_2) \quad (6.40)$$

and the Gaussian curvature

$$K = C_1 C_2. \quad (6.41)$$

For a cylinder, $H = 1/2R$ and $K = 0$. For a sphere, $H = 1/R$ and $K = 1/R^2$, For a saddle, $H = 0$ and $K = -1/R^2$.

Area of curved surface

Imagine a straight rod of length l_0 and thickness d and bend it such that it assumes the shape of a circular arc of radius R and angle $\phi = l_0/R$. In case of pure bend, the midplane is the neutral plane and remains unstretched. In a local coordinate system where the z axis points in the radial direction and $z = 0$ is the neutral plane, parts of rod at $z > 0$ are stretched and those at $z < 0$ are compressed:

$$l(z) = \phi(R + z) = \frac{l_0}{R}(R + z) = l_0 \left(1 + \frac{z}{R}\right). \quad (6.42)$$

For a rectangular plate of length l_0 and width w bent in a cylindrical shape, we have

$$A(z) = w\phi(R + z) = w\frac{l_0}{R}(R + z) = A_0 \left(1 + \frac{z}{R}\right). \quad (6.43)$$

Here $A_0 = A(z = 0) = wl_0$ is the area of the plate. We recognize that $C_1 = 1/R$ is one of the principal curvatures; the other one is 0. We conclude that

$$A(z) = A(0)(1 + C_1 z). \quad (6.44)$$

If the plate were also bent in the other direction such that it would form a cap- or a saddle-like surface, we would then have

$$A(z) = A(0) [1 + C_1 z + C_2 z] = A(0) [1 + (C_1 + C_2)z]. \quad (6.45)$$

But this is not the end of the story. Consider a sphere of radius R ; its area is $4\pi R^2$. Now increase its radius by a small amount z : The area of the larger sphere is

$$A(z) = 4\pi(R+z)^2 = 4\pi(R^2 + 2Rz + z^2) \quad (6.46)$$

$$= 4\pi R^2(1 + 2z/R + z^2/R^2) = A(0)(1 + 2z/R + z^2/R^2). \quad (6.47)$$

Knowing that for a sphere $C_1 = C_2 = 1/R$, we recognize the second term in the bracket as $(C_1 + C_2)z$. The third term is $C_1 C_2 z^2$. Thus we are led to conjecture that the complete formula for $A(z)$ for an arbitrary surface is

$$A(z) = A(0) [1 + C_1 z + C_2 z] = A(0) [1 + (C_1 + C_2)z + C_1 C_2 z^2]. \quad (6.48)$$

A mathematical aside: From Eq. (6.48) we see that the difference of the areas of two parallel surfaces $A(0)$ and $A(dz)$ that are separated by an infinitesimal dz is

$$dA = A(dz) - A(0) = A(0)(C_1 + C_2)dz \quad (6.49)$$

(We neglected the second-order term in dz .) Now $A(0)dz = dV$ is the volume of the shell between the two surfaces and so

$$\frac{\partial A}{\partial V} = C_1 + C_2 = 2H. \quad (6.50)$$

Here $H = (C_1 + C_2)/2$ is the mean curvature. — For a sphere, $A = 4\pi R^2$ and $V = 4\pi R^3/3$ so $R = (3V/4\pi)^{1/3}$ and $A = 4\pi(3V/4\pi)^{2/3}$. Thus $\partial A/\partial V = 4\pi(2/3)(3V/4\pi)^{-1/3}(3V/4\pi) = 2(3V/4\pi)^{-1/3} = 2/R = 2H$ because in a sphere, $H = R^{-1}$.

In a bilayer that is neither stretched nor compressed and is characterized by principal curvatures C_1 and C_2 measured in the midplane, the area of the neutral plane at $z = 0$, $A(0)$, is the same as the area of a flat bilayer. But the areas of the top (A) and bottom (B) surface of the bilayer are changed and this affects the free energy of the bilayer. In a symmetric bilayer with N amphiphiles in either monolayer, the average free energy per amphiphile is

$$\frac{F}{2N} = \frac{1}{2} (f_A + f_B). \quad (6.51)$$

The headgroup area of amphiphiles in the top monolayer is

$$a_A = \frac{A(z = h/2)}{N} = a(z = h/2). \quad (6.52)$$

According to Eq. (6.48), $A(z = h/2) = A(0) [1 + (C_1 + C_2)h/2 + C_1C_2h^2/4]$ and we have

$$a_A = a_0 \left[1 + (C_1 + C_2)\frac{h}{2} + C_1C_2\frac{h^2}{4} \right] \quad (6.53)$$

where $a_0 = A(0)/N$. Similarly,

$$a_B = \frac{A(z = -h/2)}{N} = a(z = -h/2) \quad (6.54)$$

and

$$a_B = a_0 \left[1 - (C_1 + C_2)\frac{h}{2} + C_1C_2\frac{h^2}{4} \right]. \quad (6.55)$$

The average energy per amphiphile in the top and bottom monolayer depends on a_A and on a_B , respectively. For simplicity, let us consider only the surface tension and the head-head repulsion:

$$f = \frac{1}{2} \left[\gamma(a_A + a_B) + W \left(\frac{1}{a_A} + \frac{1}{a_B} \right) \right]. \quad (6.56)$$

The variation of f due to curvature is given by

$$\begin{aligned} \delta f &= \frac{1}{2} \left(\gamma a_0 \left[(C_1 + C_2)\frac{h}{2} + C_1C_2\frac{h^2}{4} - (C_1 + C_2)\frac{h}{2} + C_1C_2\frac{h^2}{4} \right] \right. \\ &\quad \left. + \frac{W}{a_0} \left\{ -(C_1 + C_2)\frac{h}{2} + [(C_1 + C_2)^2 - C_1C_2]\frac{h^2}{4} \right. \right. \\ &\quad \left. \left. + (C_1 + C_2)\frac{h}{2} + [(C_1 + C_2)^2 - C_1C_2]\frac{h^2}{4} \right\} \right) \quad (6.57) \end{aligned}$$

$$= \frac{1}{2} \frac{Wh^2}{2a_0} (C_1 + C_2)^2 + \frac{1}{2} \left(\frac{\gamma a_0 h^2}{2} - \frac{Wh^2}{2a_0} \right) C_1C_2. \quad (6.58)$$

Note that in expanding a_A and a_B in the head-head repulsion, one has to keep all second-order terms:

$$\frac{1}{1 \pm uq + vq^2} \approx 1 \mp uq + (u^2 - v)q^2. \quad (6.59)$$

Bending energy*

The importance of the above result Eq. (6.58) is that it contains both a term proportional to $(C_1 + C_2)^2$ and a term proportional to $C_1 C_2$. This conclusion is not altered by including the tail-tail repulsion and the functional form of the elastic energy remains unchanged. Within a more general theory where the numbers of amphiphiles in the monolayers are not necessarily identical (so that the top and bottom monolayers contain $\chi 2N$ and $(1 - \chi)2N$ molecules, respectively), the two bending moduli are not the same and the functional form of the membrane bending energy per unit area $f_b = \delta F/[Na_0]$ is

$$f_b = \frac{1}{2}\kappa_b (C_1 + C_2)^2 + \bar{\kappa}C_1C_2 + \frac{1}{2}\lambda \left(\chi - \frac{1}{2}\right)^2 + \omega \left(\chi - \frac{1}{2}\right) (C_1 + C_2). \quad (6.60)$$

Here $\kappa_b \propto h^2$ is the bending modulus, $\bar{\kappa} \propto h^2$ is the Gaussian (saddle-splay) bending modulus, λ is the modulus measuring the energy cost of a deviation from a symmetric membrane with identical numbers of molecules in each monolayer ($\chi = 1/2$), and ω is the strength of the coupling between the mean curvature and membrane asymmetry assumed linear in both $C_1 + C_2$ and χ . This is the central result that we wanted to obtain. Both κ_b and $\bar{\kappa}$ are typically around 10^{-19} J.

Let us first estimate λ , which corresponds to the energy cost of a variation of χ at constant headgroup area in a planar bilayer. We write χ as

$$\chi = \frac{a_0}{2a_A} \quad (6.61)$$

so that for $a_A = a_0$ we have $\chi = 1/2$. Now we connect the variation of χ to the variation of a_A :

$$\delta\chi = \chi - 1/2 = -\frac{a_0}{2a_A^2}\delta a_A. \quad (6.62)$$

In equilibrium, $a_A = a_0$ and so

$$\delta\chi = -\frac{\delta a_A}{2a_0}. \quad (6.63)$$

This gives

$$\frac{\delta F}{Na_0} = \frac{1}{2}\lambda \left(\chi - \frac{1}{2}\right)^2 = \frac{1}{2}\lambda \frac{1}{4} \left(\frac{\delta a}{a_0}\right)^2 = \frac{1}{2}\kappa_A \left(\frac{\delta a}{a_0}\right)^2 \quad (6.64)$$

In the last equality, we have used the area compressibility formula [Eq. (6.38)]. Thus $\lambda = 4\kappa_A$.

Exchange of lipid molecules between monolayers

If the headgroups of the two monolayers are polymerized, then they cannot jump from one monolayer to the other as a response of the bilayer to stress due to curvature. This is the limit of blocked exchange. In the opposite limit, exchange is possible and in this case a finite $C_1 + C_2$ will induce a deviation of χ from its equilibrium value of $\chi = 1/2$ in a flat membrane. To first order, we can assume that $\chi - 1/2$ is proportional to $C_1 + C_2$:

$$\chi - \frac{1}{2} = \eta d(C_1 + C_2). \quad (6.65)$$

η is the coupling parameter and d has been inserted to make η dimensionless.

Using this approximation, we can rewrite the elastic energy density [Eq. (6.60)] as

$$\begin{aligned} f_b &= \frac{1}{2}\kappa_b(C_1 + C_2)^2 + \bar{\kappa}C_1C_2 + \frac{1}{2}\lambda\eta^2d^2(C_1 + C_2)^2 + \omega\eta d(C_1 + C_2)^2 \\ &= \frac{1}{2}(\kappa_b + \lambda\eta^2d^2 + 2\omega\eta d)(C_1 + C_2)^2 + \bar{\kappa}C_1C_2 \\ &= \frac{1}{2}\kappa(C_1 + C_2)^2 + \bar{\kappa}C_1C_2 \end{aligned} \quad (6.66)$$

In the last line, we have introduced the rescaled bending modulus

$$\kappa = \kappa_b + \lambda\eta^2d^2 + 2\omega\eta d. \quad (6.67)$$

In case of **free exchange** between monolayers, any bending is compensated for by an adjusted χ so as to minimize the bending modulus. By differentiating κ with respect to η we find that $\eta_{min} = -\omega/\lambda d$ and

$$\kappa_{min} = \kappa_b - \frac{\omega^2}{\lambda} \quad (6.68)$$

The opposite extreme is **blocked exchange**. In this case, there is no coupling between χ and C_+ and $\eta = 0$ so that $\kappa = \kappa_b$.

Spontaneous curvature

For some membranes, a planar bilayer is not the equilibrium state. To model bilayers where the minimal elastic energy corresponds to a state with $C_+ \neq 0$, we extend Eq. (6.66) by a term of the form $-\kappa C_0 C_+$. C_0 is a phenomenological parameter. So we have

$$f_b = \frac{1}{2}\kappa(C_1 + C_2)^2 + \bar{\kappa}C_1C_2 - \kappa C_0(C_1 + C_2) \quad (6.69)$$

$$= \frac{1}{2}\kappa(C_+ - C_0)^2 + \bar{\kappa}C_1C_2 - \frac{1}{2}\kappa C_0^2. \quad (6.70)$$

(In passing to the last equality, we merely completed the square.) This is the so-called Helfrich free energy.

6.3 Vesicles*

Phospholipid bilayer membranes are the basic building block of animal cells, and lipid vesicles can also be formed in an artificial environment. Their size can vary from a few 10 nm to 10 μm ; in any case, the bilayer membrane thickness of a few nm is considerably smaller than the vesicle diameter. The total elastic free energy of a vesicle is the integral of f_b over the whole membrane area:

$$F = \frac{\kappa}{2} \oint (C_+ - C_0)^2 dA + \bar{\kappa} \oint C_1 C_2 dA. \quad (6.71)$$

As the membrane is very thin compared to vesicle size, we can integrate over the neutral surface.

In principle, we should also include the membrane stretching energy of the form

$$\frac{\kappa_A (A - A_0)^2}{2 A_0} \quad (6.72)$$

where A_0 is the reference area of the vesicle. However, in large vesicles the typical magnitude of this term for any nonzero variation of A would be much larger than the bending terms. For example, in a vesicle of radius 10 μm the relative area variation such that the stretching energy would be comparable to the bending energy is minute:

$$\frac{A - A_0}{A_0} = \sqrt{\frac{\kappa}{\kappa_A A_0}} \sim 10^{-5}. \quad (6.73)$$

We have used $\kappa \sim 10^{-19}$ J and $\kappa_A \sim 0.1$ J/m². So we usually assume that the membrane is unstretchable and that its area is constant.

The other common approximation is that of a constant volume. To some extent, the membrane itself is permeable to water. But usually the water is not pure and any solvent present gives rise to an osmotic pressure. So any exchange of water that may take place between the vesicle interior and the environment would inevitably destroy the pressure balance and the ensuing pV terms would be much larger than the bending energy. As a result, vesicle volume is essentially fixed.

Thus the vesicle elastic energy [Eq. (6.71)] is minimized subject to two constraints: $A = \text{const.}$ and $V = \text{const.}$ Moreover, the Gaussian elasticity

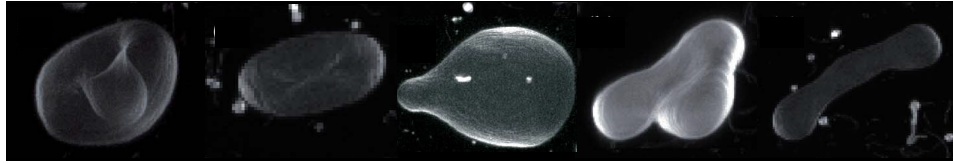


Figure 6.6: A few typical vesicle shapes: Stomatocytes, discocytes, pears, triangular non-axisymmetric shape, and cigar. Adapted from Ref. [27].

term is irrelevant as long as we study vesicles of fixed topology. The Gauss-Bonnet theorem says that the integral of the Gaussian curvature over a closed area is equal to $4\pi(1 - g)$ where g is the genus (the number of holes). In a vesicle of spherical topology, $g = 0$ and we can easily verify the above result: For a sphere, $C_1 = C_2 = 1/R$ and $\oint C_1 C_2 dA = 4\pi$. (In a torus, $g = 1$ and in a two-hole button $g = 2$.) Because of the Gauss-Bonnet theorem, one needs not to worry about the Gaussian term as long as the topology of the vesicle is fixed.

One of the key parameters that determine the shape of a vesicle is its reduced volume. The reduced volume is defined as the ratio of the vesicle volume (V) and the volume of a sphere whose area is the same as the area of the vesicle in question (A), and is defined by

$$v = \frac{6\sqrt{\pi}V}{A^{3/2}}. \quad (6.74)$$

In a sphere, $v = 1$ and in all other shapes $v < 1$. v can be considered a measure of vesicle inflatedness. In the spontaneous-curvature model, the other parameter specifying vesicle shape is the spontaneous curvature; in the bilayer-couple and the area-difference-elasticity model, it is the monolayer area difference and the preferred monolayer area difference

$$\Delta A = h \oint (C_1 + C_2) dA = 2d \oint H dA \quad (6.75)$$

proportional to the integrated mean curvature of the vesicle. In all three models, the classes of vesicle shapes $g = 0$ include ellipsoids, discocytes, stomatocytes, prolate shapes (cigars), pears, nonaxisymmetric shapes (e.g., starfish-like shapes), necklaces, etc. (Fig. 6.6).

Homework problems

1. Third virial coefficient for hard spheres, Carnahan-Starling equation of state (D. Grošelj)
2. Thermodynamic inconsistency (J. Zmrzlikar)
3. Diagrammatic expansion of pair functions (G. Posnjak)
4. Virial expansion of equation of state
5. Derivation of Frank free energy (B. Kavčič)
6. Frederiks transition (M. Krajnc)
7. Landau-de Gennes theory of nematic-isotropic transition (A. Horvat)
8. Undulation instability in smectics (N. Štorgel)
9. Polymer demonstration experiments (M. Medenjak, I. Kukuljan)
10. Worm-like chain (J. Zobec)
11. Polyelectrolytes (J. Mur)
12. Rotational isomeric states (M. Trček)
13. Derjaguin approximation (M. Čančula)
14. Depletion interaction between spheres (N. Rosenstein)
15. Self-consistent field theory of polymers (N. Adžić)
16. Classical approximation for polymer brushes (T. Mohorič)
17. Lattice models of polymer chains (B. Jenčič)
18. Elasticity of networks (J. Sablič)

19. Phase diagram of binary solution (Ž. Kos)
20. Smectic-superconductor analogy (B. Mavrič)
21. Small-angle domain walls in smectics (T. Suhovršnik)
22. Limiting shapes of bilayer vesicles (M. Zadnik)
23. Elasticity of 2D networks (J. Sablič)
24. Membrane disks and cups (N. Lopič)
25. Persistence length of amphiphilic membranes (T. Dobravec)
26. Fractals in soft matter (K. Vozel)
27. Diffusion- and reaction-limited aggregation (R. Grah)
28. Complex amphiphile/diblock copolymer morphologies (A. Marin)
29. Minimal surfaces in soft matter (M. Vitek)
30. Polymer chain in a cylindrical tube (T. Verbovšek)
31. Nematic liquid crystal at a grooved wall (E. Ule)
32. Escaped nematic structure in a capillary (T. Parkelj)
33. (U. Tomšič)

Bibliography

- [1] R. A. L. Jones, *Soft Condensed Matter* (Oxford University Press, Oxford, 2002).
- [2] G. Strobl, *The Physics of Polymers* (Springer, Berlin, 1997).
- [3] J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic Press, San Diego, 1986).
- [4] C. Caccamo, Phys. Rep. **274**, 1 (1996).
- [5] J. A. Barker and D. Henderson, Rev. Mod. Phys. **48**, 587 (1976).
- [6] P. Bolhuis and D. Frenkel, J. Chem. Phys. **106**, 666 (1997).
- [7] T. A. Witten, *Structured Fluids* (Oxford University Press, Oxford, 2004).
- [8] M. Daoud and C. E. Williams (eds.), *Soft Matter Physics* (Springer, Berlin, 1999).
- [9] I. W. Hamley, *Introduction to Soft Matter* (Wiley, Chichester, 2000).
- [10] J. W. Goodby, Curr. Opin. Colloid Interf. Sci, **7**, 326 (2002).
- [11] D. R. Nelson, T. Piran, and S. Weinberg (eds.), *Statistical Mechanics of Membranes and Surfaces* (World Scientific, Singapore, 2004).
- [12] S. Hyde, S. Andersson, K. Larsson, Z. Blum, T. Landh, S. Lidin, and B. W. Ninham, *The Language of Shape* (Elsevier, Amsterdam, 1997).
- [13] J. Duran, *Sands, Powders, and Grains* (Springer, Berlin, 1999).
- [14] P. J. Collings and M. Hird, *Introduction to Liquid Crystals* (Taylor & Francis, London, 1997).

- [15] P.-G. de Gennes and J. Prost, *The Physics of Liquid Crystals*, (Clarendon Press, Oxford, 1993).
- [16] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, 1995).
- [17] T. Kawakatsu, *Statistical Physics of Polymers* (Springer, Berlin, 2001).
- [18] P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, 1979).
- [19] D. Stauffer and A. Aharony, *Introduction To Percolation Theory* (Taylor & Francis, London, 1985).
- [20] J. N. Israelachvili, *Intermolecular and Surface Forces* 3rd ed. (Academic Press, Waltham, 2011).
- [21] W. B. Russel, D. A. Saville, and W. R. Schowalter, *Colloidal Dispersions* (Cambridge University Press, Cambridge, 1992).
- [22] D. F. Evans and H. Wennerström, *The Colloidal Domain* (Wiley-VCH, New York, 1999).
- [23] K.-H. Lin, J. C. Crocker, V. Prasad, A. Schofield, D. A. Weitz, T. C. Lubensky, and A. G. Yodh, *Phys. Rev. Lett.* **85**, 1770 (2000).
- [24] V. A. Markel, V. M. Shalaev, E. B. Stechel, W. Kim, and R. L. Armstrong, *Phys. Rev. B* **53**, 2425 (1996).
- [25] A. Ben-Shaul and W. M. Gelbart, *Statistical Thermodynamics of Amphiphile Self-Assembly: Structure and Phase Transitions in Micellar Solutions*, in *Micelles, Membranes, Microemulsions, and Monolayers*, W. M. Gelbart, A. Ben-Shaul, and D. Roux (eds.) (Springer, Berlin, 1994).
- [26] A. Ben-Shaul, *Molecular Theory of Chain Packing, Elasticity and Lipid-Protein Interaction in Lipid Bilayers*, in *Structure and Dynamics of Membranes*, R. Lipowsky and E. Sackmann (eds.) (North-Holland, Amsterdam, 1995).
- [27] A. Sakashita, N. Urakami, P. Ziherl, and M. Imai, *Soft Matter* **8**, 8569 (2012).