Advanced Statistical Mechanics

K. Kroy

Leipzig, 2017*

Contents

---

*The script is not meant to be a substitute for reading proper textbooks nor for dissemination. (See the notes for the introductory course for background information.) Comments and suggestions are highly welcome.

> Die Begründung der statistischen Mechanik bereitet begriffliche Schwierigkeiten. Die dabei durchzuführenden Rechnungen sind jedoch meist einfach. Dieser Sachverhalt kehrt sich bei den Anwendungen der statistischen Mechanik gerade um. Begrifflich gibt es dabei nahezu keine, dafür um so mehr mathematische Schwierigkeiten.
>
> W. Brenig

1 The empty canvas

The aim of this part of the lecture is to introduce general theoretical tools that have been developed and used in many areas of physics to address interacting many-body systems on a microscopic basis.

Statistical Mechanics provides a microscopic foundation of the thermostatics of many-body systems, i.e., it answers the question how the properties of the few degrees of freedom conventionally used to describe macroscopic material behavior emerges from the many degrees of freedom of all contributing atoms. The central paradigmatic many-body system discussed in introductory courses is an isothermal (classical) ideal gas. This toy model with no interactions and no packing structure is briefly recapitulated to set the stage. Thermodynamically, the ideal gas is defined in an operational way by two equations of state: (i) the caloric equation or energy equation

$$U/N \equiv u = 3k_BT/2\,, \tag{1}$$

which can be measured in a calorimeter (it amounts to equipartition of energy U among dN degrees of freedom in $d = 3$ space dimensions and contains the full statistical information of the model); (ii) an equation linking energy to mechanical properties (e.g., the pressure p),

$$pV/N \equiv pv = 2u/3 = k_BT\,. \tag{2}$$

The first equality contains new information, namely about the kinematics of massive Newtonian particles, the second simply substitutes the "heat quantum" k_BT from Eq. (1). Using Eqs. (1) and (2) to express the intensive variables in terms of the (densities of) extensive ones, and invoking homogeneity, the differential fundamental relation (per particle)

$$ds(u,v) = \frac{1}{T}du + \frac{p}{T}dv \tag{3}$$

in a 3-dimensional thermodynamic state space spanned by energy U, volume V, and particle number N can be integrated to yield the fundamental relation $S(U,V,N) = Ns(U/N, V/N)$, from which all thermostatic properties follow.

In Statistical Mechanics, which represents a system in terms of all of its microscopic degrees of freedom, one is used to addressing the fundamental relation directly—at least on a conceptual level—via the (e.g. canonical) partition sum

$$Z \equiv \int \mathrm{d}\Gamma \, e^{-\beta H} = e^{-\beta F(T,V,N)} \qquad \mathrm{d}\Gamma \equiv \frac{\Pi_i \mathrm{d}\mathbf{q}_i \mathrm{d}\mathbf{p}_i}{N! h^{dN}} \ . \tag{4}$$

Integrating out the microscopic degrees of freedom at fixed thermodynamic parameters is called coarse graining. It transforms the the Hamiltonian into a free energy (a kind of Hamiltonian with much less variables). This is how it works in principle but rarely in practice. In fact, most of the following serves to introduce more practical approaches to avoid doing this monstrous integral, since a brute-force approach works only for toy-models, such as the ideal gas. In fact, with the Hamiltonian $H = \sum_i \mathbf{p}_i^2/2m$, one easily finds the fundamental relation

$$Z = (V/\lambda_T^3)^N/N! \quad \Rightarrow \quad F = -k_BT\ln Z = Nk_BT[\ln(n\lambda_T^3) - 1] \tag{5}$$

The *thermal wavelength*

$$\lambda_T \equiv \frac{h}{\sqrt{2\pi m k_B T}} \tag{6}$$

appears as a natural length scale in a quantum statistical description, and defines an elementary unit cell size in the classical configuration space. It is the de Broglie wavelength of a typical gas particle of mass m with an energy of about k_BT, as dictated by equipartition. The naive interpretation of V/λ_T^3 as the number of "available classical states" for each of N non-interacting particles turns out to be inappropriate if these are indistinguishable, though. Since particle permutations produce no new states in this case, each particle can only explore the specific volume $v = 1/n$ instead of V, and there are only $\mathcal{W} \approx v/\lambda_T^3$ classical states per particle available, which is also apparent from the entropy $S = (U - F)/T = (U + pV - \mu N)/T$, which is essentially (up to a thermodynamically irrelevant gauge constant) the chemical potential μ, namely

$$s/k_B = 5/2 - \mu/k_BT \ , \qquad \mu/k_BT = \ln(n\lambda_T^3) \ . \tag{7}$$

It essentially counts the effective degrees of freedom that participate in energy partition according to Eq. (1), in units of Boltzmann's constant k_B, evidencing that individual-particle degrees of freedom are not independent for indistinguishable particles, and pointing at an entropy crisis or "freezing" at low T.

Not too surprisingly, the model therefore breaks down if one attempts to raise the density or to lower the temperature beyond the limit $n\lambda_T^3 \simeq 1$. Equipartition, saying that the same energy is stored in each degree of freedom (in particular also in the thermodynamically unresolved ones, which constitute sort of a black box constantly absorbing other forms of energy), is a classical approximation. Quantum mechanics shows that degrees of freedom freeze out at low temperature T, causing the condensation of an increasing fraction of particles into their quantum ground states. These particles do not scatter any more between different states. Therefore they carry no heat or entropy and behave very differently from what is expected for a gas. The number of particles available to contribute to the classical gas behavior, or to any sort of thermodynamics, is thus reduced and the thermodynamic properties are modified with respect to the classical expectation; e.g., it is found that the pressure is bound from below for fermions and from above for bosons. With a grain of salt, one can say that the thermodynamics at low T is essentially all due the normal fraction of the gas, which has a reduced density $n_c = n(T/T_c)^\zeta$ with a characteristic exponent ζ and BEC/Fermi-temperature T_c, but otherwise still behaves like a classical ideal gas.

2 Packing structure and material behavior

2.1 Pair interactions and pair correlations

Pair interactions

For atoms and molecules, potential interactions can usually not be neglected. It is usually via these “direct interactions” (as opposed to the “statistical” exchange interactions encoded in the chemical potential μ of an ideal gas) that we notice the existence of these particles, in the first place. And the existence of various phases of the same substance and its physical properties are primarily the result of strong mutual particle interactions. The reason that the model of an ideal (interaction-free) gas is useful at all, and in fact highly successful, is due to the fact that many elementary excitations in condensed matter may (after quantization) to a good approximation be described as dilute gases of quasi-particles (so-called “phonons”, “magnons”, “electrons”, “holes”, etc.), at least at low temperatures. Interactions govern the phase transitions into condensed states, their packing structure and energetics, and, last but not least, the often complex and composite (hybrid) nature of the elementary excitations that may then, eventually, be treated as dilute Bose or Fermi gases.

Before entering the discussion of direct interactions, it is useful to estimate their importance relative to the effective “statistical interactions” arising from the exchange statistics of indistinguishable particles. Roughly, the latter should start to matter once the effective statistical pair potential becomes of similar range and strength as the direct interactions. A common feature of all direct interactions between atoms and molecules is their strong hard-core repulsion at distances on the order of a few Bohr radii $a_B = \hbar^2/m_e e^2$. For the weak statistical interactions to become relevant, they need to be of considerably longer range,

$$\lambda_T/a_B \gg 1 \qquad \Rightarrow \qquad k_BT \ll \mathrm{Ry}\, m_e/m = \mathcal{O}(\mathrm{m\,eV}) \ , \tag{8}$$

where m_e and m denote the masses of the electron and the atom, respectively. Also note that at the characteristic distance λ_T the statistical interaction strength is only about $\pm 2 \cdot 10^{-3} k_BT$. In contrast, the strength of the repulsion is proportional to the electron density, which decays exponentially and is down to about 1 Ry$/e$ at the nominal atom radius. Altogether, this argument thus suggests that atoms and molecules have to be cooled down substantially below room temperature, if not to the lowest temperatures in the universe, before the exchange interactions become relevant. A related estimate via the comparison of the direct and exchange contributions to the second virial coefficient is discussed in the exercises. Both estimates show that the energetics and packing structure of atoms and molecules is usually completely dominated by "direct" potential interactions[1]. Of course, these direct interactions themselves arise as an effective picture from a complex combination of the wave-like nature of matter (uncertainty relation), exchange interactions (Pauli principle), and Coulomb interactions within the electronic degrees of freedom of the atoms, as discussed in atomic physics.

The most convenient starting point for studying the effects due to *direct interactions* on the many-body physics is the approximation of pure pair interactions, i.e., pair-wise additive interactions between the individual particles. In the following, *only* such pair interactions $\mathcal{V}(\{\mathbf{r}_j\}) = \sum_{i<j} \nu(\mathbf{r}_i - \mathbf{r}_j)$ in an Euclidean configuration space, corresponding to the Hamiltonian

$$H = \sum_i \frac{\mathbf{p}_i^2}{2m} + \sum_{i<j} \nu(\mathbf{r}_i - \mathbf{r}_j) \ , \tag{9}$$

are considered. This excludes the effective interactions due to the exchange statistics and other possible M-particle interactions with $M > 2$ from the very beginning. And even without invoking quantum mechanics there are many situations where this is inappropriate. As an example one may mention the effective depletion attractions induced between two tagged hard spheres by other hard spheres that cannot enter the depletion zones around these particles (sketch). These effective attractions are, in general, not pair-wise additive[2] although they arise from the purely pairwise interactions between the individual hard-spheres. This example therefore also serves to illustrate the point that the restriction to "elementary" pair interactions does of course not exclude the possibility that complicated many-body correlations arise and govern the packing structure and macroscopic behavior.

Now, the general task is to calculate the partition sum, or rather (since the integration over momenta can immediately be performed) the configuration integral

---

[1] For conduction electrons in solids exchange interactions are crucial, at room temperature.

[2] A dilute, strongly bi-disperse mixture of hard spheres provides an exception (exercises).

over the coordinates $\mathbf{q}_i$ for the Hamiltonian (9). Except for a few toy models, this is analytically completely out of reach, which leaves one with essentially two choices. The first one is to calculate Z numerically. This replaces an arbitrary generic problem in equilibrium statistical mechanics by a rather specific technical challenge, namely the accurate numerical evaluation of an extremely high-dimensional integral. The universal method to address this is called *Monte Carlo simulation*. To get the basic idea, imagine a simple 2-dimensional example, namely the evaluation of the area of a circle. Draw your circle inside a sand pit of known area and let the kids randomly throw pebbles into the sand pit. At the end of the day you collect the pebbles from inside and outside the circle area and compare their accumulated weights to get an estimate of the ratio of the circle area to the known sand pit area (see exercises). An alternative approach (less suitable for the exploitation of child labor) is pursued in the following.

Pair correlations in real space: the radial distribution function

As another restriction, the following focuses almost exclusively on the discussion of pair correlations. They contain in many cases the most salient information. Higher-order correlations may well contain essential additional information, but they are certainly more difficult to measure and to calculate. In fact, a framework based on pair-interactions and pair-correlations is to date the most common and most practical way to approach many of the formidable challenges provided by many-body systems. But beware! In general, pair correlations will not only not tell us the whole story, but they may even fail to provide a good qualitative picture of the overall behavior. Concentrated hard spheres can serve as a striking example. They interact only pairwise, yet exhibit a dramatic slowing down of their dynamics, called a glass transition, without any obvious sign in their pair correlations[3]. Similarly, complex ring-like bridges involving subtle correlations between more than just pairs of particles are held responsible for the jamming of granular flows (which complicates the construction of a good hourglass).

The microscopic particle concentration is

$$\hat{n}(\mathbf{r}) \equiv \sum_i^N \delta(\mathbf{r} - \mathbf{r}_i) \ . \tag{10}$$

The ensemble-averaged position-dependent “1-body” or “1-point” density (in the grand-canonical formulation)

$$n(\mathbf{r}) = \langle \hat{n}(\mathbf{r}) \rangle = \langle \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \rangle = \langle N \delta(\mathbf{r} - \mathbf{r}_1) \rangle \tag{11}$$

simply contains the information how many particles are found on average in a certain region of configuration space. For a homogeneous (i.e., translation-invariant)

---

[3] In contrast, their crystalization is well detectable in the pair correlations.

system, $n(\mathbf{r}) = n \equiv N/V$ it is just the overall density, which is constant. The more interesting function to look at is then the 2-point density or *pair correlation function*. It shall here be defined as $\langle \hat{n}(\mathbf{r})\hat{n}(\mathbf{r}')\rangle$ minus the self-correlations[4],

$$n(\mathbf{r}', \mathbf{r}'') \equiv \langle \sum_{i \neq j} \delta(\mathbf{r}' - \mathbf{r}_i)\delta(\mathbf{r}'' - \mathbf{r}_j)\rangle = \langle N(N-1)\delta(\mathbf{r}' - \mathbf{r}_1)\delta(\mathbf{r}'' - \mathbf{r}_2)\rangle \ . \tag{12}$$

There are usually still interesting pair correlations if the system is translation invariant. Then $n(\mathbf{r}', \mathbf{r}'')$ is solely dependent on the *relative* position $\mathbf{r} = \mathbf{r}' - \mathbf{r}''$. This can be exploited by setting the origin of the coordinate system at the center of an arbitrary particle and defining a dimensionless correlation function $g(\mathbf{r})$ by

$$\begin{aligned} n^2 g(\mathbf{r}) \equiv& \frac{1}{V} \int \mathrm{d}\mathbf{r}' \, n(\mathbf{r} + \mathbf{r}', \mathbf{r}') \\ n g(\mathbf{r}) =& \frac{1}{\langle N\rangle} \int \mathrm{d}\mathbf{r}' \, n(\mathbf{r} + \mathbf{r}', \mathbf{r}') = \frac{1}{\langle N\rangle} \int \mathrm{d}\mathbf{r}' \, \langle \sum_{i \neq j} \delta(\mathbf{r} + \mathbf{r}' - \mathbf{r}_i)\delta(\mathbf{r}' - \mathbf{r}_j)\rangle \\ =& \frac{1}{\langle N\rangle} \langle \sum_{i \neq j} \delta(\mathbf{r} + \mathbf{r}_j - \mathbf{r}_i)\rangle = \frac{1}{\langle N\rangle} \langle N(N-1)\delta(\mathbf{r} + \mathbf{r}_1 - \mathbf{r}_2)\rangle \ ; \end{aligned} \tag{13}$$

$$g(\mathbf{r}) \overset{\mathbf{r}_{12} \equiv \mathbf{r}_2 - \mathbf{r}_1}{=} \frac{V}{\langle N\rangle^2} \langle N(N-1)\delta(\mathbf{r} - \mathbf{r}_{12})\rangle \overset{N = \text{const.}}{=} V \langle \delta(\mathbf{r} - \mathbf{r}_{12})\rangle \tag{14}$$

This manifestly translation invariant form of the pair correlation function encodes the neighbor correlations in the simplest possible way, namely as a local particle density as seen from the position of an arbitrary particle. If the system is moreover isotropic, the argument of $g(\mathbf{r})$ is actually the distance $r \equiv |\mathbf{r}|$, and one speaks of the *radial distribution function* $g(r)$. It gives the probability density $4\pi r^2 g(r)/V$ to find a particle at distance r from an arbitrarily chosen test particle. For a statistically homogeneous and isotropic system the full information about the packing structure, as far as pairs of particles are concerned (i.e. excluding the information, with which probability three or more particles are found in a certain correlated state), is thus encoded in a single scalar function of a scalar variable. For hard core particles[5] $g(r \to 0) = 0$. For systems without long-range order, such as liquids and gases, the limit $g(r \to \infty) = 1$ exists, for crystals $g(r)$ will oscillate indefinitely around it. For Fourier transforming $g(\mathbf{r})$, it is customary to subtract the 1, which would otherwise always produce a δ-function corresponding to the forward scattering of a homogeneous material. The resulting function, which only retains the non-trivial pair correlations or "pair fluctuations", is often denoted by

$$h(\mathbf{r}) \equiv g(\mathbf{r}) - 1 = \left[\langle \hat{n}(\mathbf{r})\hat{n}(0)\rangle - \langle \hat{n}(0)\rangle^2\right] / n^2 \ . \tag{15}$$

(The last expression should be understood as a notational convention; the self correlations drop out.)

---

[4] The experts often tacitly use the same notation, no matter whether they want the self-correlations in or out, which has then to be decided from the context.

[5] Remember that self-correlations, which would contribute a $\delta(r)$ to $ng(r)$, were subtracted.

Pair correlations in Fourier space: the structure factor

It should be familiar from elementary optics that the Fraunhofer interference patterns resulting from scattering in optically dilute (i.e. essentially transparent) media correspond to a Fourier transform of the pattern of point scatterers. The relevant non-technical information, namely the (relative) scattering intensity with scattering vector $\mathbf{q}$, is contained in the *structure factor*

$$\begin{aligned} S_{\mathbf{q}} \equiv & \frac{1}{\langle N\rangle}\langle \sum_{ij} \exp[-i\mathbf{q}\cdot(\mathbf{r}_i-\mathbf{r}_j)]\rangle \\ =& \frac{1}{\langle N\rangle}\int \mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r}''\, \exp[-i\mathbf{q}\cdot(\mathbf{r}'-\mathbf{r}'')]\langle \hat{n}(\mathbf{r}')\hat{n}(\mathbf{r}'')\rangle \\ =& 1+\frac{1}{\langle N\rangle}\int \mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r}''\, \exp[-i\mathbf{q}\cdot(\mathbf{r}'-\mathbf{r}'')]n(\mathbf{r}',\mathbf{r}'') \\ =& 1+n\int \mathrm{d}\mathbf{r}\, \exp(-i\mathbf{q}\cdot\mathbf{r})g(\mathbf{r}) \\ =& 1+(2\pi)^d n\delta(\mathbf{q})+n\int \mathrm{d}\mathbf{r}\, \exp(-i\mathbf{q}\cdot\mathbf{r})h(\mathbf{r})\ . \end{aligned} \tag{16}$$

For the first equality, introduce a $\mathbf{1}$ in the form $\int \mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r}''\, \delta(\mathbf{r}'-\mathbf{r}_i)\delta(\mathbf{r}''-\mathbf{r}_j)$. The leading 1 emerging in the third line corresponds to the incoherent scattering from $\langle N\rangle$ independent particles and results from the self-correlations in the density auto-correlation $\langle \hat{n}(\mathbf{r})\hat{n}(\mathbf{r}')\rangle$. The remaining terms are due to interference and encode correlations. The δ-function amounts to the forward "scattering" from the homogeneous contribution contained in $g(\mathbf{r})$ and subtracted in $h(\mathbf{r})$. It is quite often tacitly omitted, which makes the limit $q\to 0$ of the structure factor continuous, in agreement with the experimental procedure of using a beam stop to block the contribution at $q=0$. The last term contains the interesting information in the form of a Fourier transform of the pair fluctuations. The information encoded in small-angle scattering ($q\neq 0$) is the integral over the fluctuations $h(\mathbf{r})$, hence according to Eqs. (14), (15)

$$S_{\mathbf{q}\to\mathbf{0}} \sim 1+n\int \mathrm{d}\mathbf{r}\, h(\mathbf{r}) = 1+\langle N(N-1)\rangle/\langle N\rangle - \langle N\rangle\,. \tag{17}$$

Using a result from introductory statistical mechanics relating number fluctuations and compressibility κ_T (exercises), we find the so-called *compressibility sum rule*

$$S_{\mathbf{q}\to\mathbf{0}} \sim \langle (N-\langle N\rangle)^2\rangle/\langle N\rangle = nk_BT\kappa_T\ . \tag{18}$$

It provides a reverse perspective at the relation between fluctuations and response. While fluctuations are usually understood to increase in magnitude because of a weakening of the restoring forces, which are controlled by the response coefficient, Eq. (18) expresses the value of a response coefficient by an integral over the fluctuations. It thereby clearly displays the connection between a diverging correlation length, which entails a (spatially) slow decay of the fluctuations encoded in $h(\mathbf{r})$, and a divergence of response coefficients at a critical point.

In summary, the pair distribution completely determines one of the most important experimental observables of the packing structure of a material, namely the static structure factor. Its spatial integral moreover also determines the materials' mechanical strength (its compression modulus) via a sum rule, which provides a special case of the general fluctuation-response theorem, one of the major predictions by statistical mechanics that repeatedly reappears in various generalizations throughout this part of the lecture.

Exercise: Argue loosely that for dilute gases, one can calculate $g(r)$ by only considering two particles, and that, in this limit, $g(\mathbf{r}) \sim e^{-\beta\mathcal{V}(\mathbf{r})}$, where $\mathcal{V}(\mathbf{r})$ represents interaction potential between the particles. Use this to show that $g(\mathbf{r}) = \theta(r-\sigma)$ for a gas of hard spheres with diameter σ and density $n \ll \sigma^{-3}$. Calculate and sketch the structure factor S_q and discuss the small angle scattering. Sketch your qualitative expectations for $g(r)$ and S_q for a liquid, for a crystal, and close to a liquid-vapor spinodal.

2.2 Correlation functions and constitutive relations

Exploiting the simplifications ensuing from the limitation to pair interactions, Eq. (18) can be taken somewhat further, to express the two constitutive equations (or equations of state) of an interacting gas in terms of the pair distribution function $g(\mathbf{r})$.

The caloric equation of state or *energy equation*, Eq. (1) is readily generalized,

$$U = \langle H \rangle = 3\langle N \rangle k_B T/2 + \langle \mathcal{V} \rangle \ . \tag{19}$$

Noting that all pairs of particles give the same contribution to the last term, it may be rewritten as

$$\langle \mathcal{V} \rangle = \langle N(N-1)\nu(\mathbf{r}_{12}) \rangle/2 \ . \tag{20}$$

Now, introducing a $\mathbf{1}$ in the form $\int \mathrm{d}\mathbf{r}\, \delta(\mathbf{r}-\mathbf{r}_{12})$ and using Eq. (14), this may immediately be rephrased in the form

$$U = \langle H \rangle = \frac{3}{2}\langle N \rangle k_B T \left[1 + \frac{n}{3} \int \mathrm{d}\mathbf{r}\, g(\mathbf{r}) \beta\nu(\mathbf{r}) \right] \ . \tag{21}$$

Similarly, the thermal (mechanical) equation of state may be extended and related to the radial distribution,

$$\beta p = -\beta \partial_V F|_T = \partial_V \ln\bigl(Z_{\mathrm{ig}} Q\bigr) = n + Q^{-1}\partial_V Q \ . \tag{22}$$

The normalized configuration integral Q is defined via $Z \equiv Z_{\mathrm{ig}} Q$, hence (in the canonical ensemble)

$$Q = \int \frac{\mathrm{d}^N \mathbf{r}}{V^N}\, e^{-\beta\mathcal{V}(\{\mathbf{r}_j\})} \ . \tag{23}$$

Due to the normalization the only volume dependence in Q is due to $\nu(\mathbf{r}_{ij})$, i.e. $\partial_V \nu(\mathbf{r}_{ij}) = \nabla\nu(\mathbf{r}_{ij})\cdot\partial_V \mathbf{r}_{ij}$. Exploiting homogeneity and writing $\mathbf{r}_{ij} = V^{1/3}\xi_{ij}$ with volume-independent matrix elements ξ_{ij}, one finds $\partial_V \mathbf{r}_{ij} = \mathbf{r}_{ij}/3V$ and

$$\partial_V Q = -\int \frac{\mathrm{d}^N\mathbf{r}}{V^N}\,\frac{N(N-1)}{2}\frac{\beta\,\mathbf{r}_{12}}{3V}\cdot\nabla\nu(\mathbf{r}_{12})e^{-\beta\mathcal{V}(\{\mathbf{r}_j\})}\ . \tag{24}$$

Using again the trick with the inserted delta function, and keeping the N's inside the averages to allow for a grand-canonical interpretation if required, one thus has

$$Q^{-1}\partial_V Q = -\frac{1}{6k_BTV}\int \mathrm{d}\mathbf{r}\,\mathbf{r}\cdot\nabla\nu(\mathbf{r})\langle N(N-1)\delta(\mathbf{r}-\mathbf{r}_{12})\rangle \tag{25}$$

With Eq. (14) the average can be rewritten as $\langle N\rangle^2 g(\mathbf{r})/V$, so that the correction to the thermo-mechanical constitutive equation of the ideal gas resulting from the potential interactions is found to be given by the average of the virial (weighted by the pair distribution), hence the name *virial equation of state*

$$p = nk_BT\left[1-\frac{n}{6k_BT}\int \mathrm{d}\mathbf{r}\,g(\mathbf{r})\mathbf{r}\cdot\nabla\nu(\mathbf{r})\right]\ . \tag{26}$$

Equations (18), (21) and (26) show how *thermodynamics follows from packing structure*[6]. They reduce the task of deriving the thermodynamic equations of state to integrations over the spatial density fluctuations. Or, more precisely, the thermodynamics of a homogeneous isotropic fluid with pair interactions can be expressed solely in terms of the two-point correlation function that reduces to $g(\mathbf{r})$ for a homogeneous system. To make any practical use of this formal result, one has to find $g(\mathbf{r})$ in a continuous parameter region, which is indeed the central task of so-called *liquid-state theories* and of many Monte Carlo simulations. In particular, a very educated guess about the general form of $g(\mathbf{r})$ and $\nu(\mathbf{r})$ is indispensable to apply the result in reverse, namely to infer aspects of the microscopic packing structure and interaction potential from macroscopic thermodynamic measurements. Historically, we owe much of our knowledge about molecular interactions to this connection, and thanks to recent developments in microscopy, it might become a fruitful path to explore even further, in the future. An important advantage of such an approach (over attempts to address the partition sum) is that it readily generalizes to conditions far from equilibrium, where the usual partition sum and free energies make no sense.

2.3 Dilute gases and the virial expansion

It is intuitively obvious that encounters between particles in a dilute gas are rare. Then, multiple collisions are probably very weakly correlated and therefore

---

[6] Three constitutive equations cannot be independent, of course, and errors of approximate liquid-state theories are sometimes estimated (or reduced) by comparing (or superimposing) the predictions obtained from Eq. (26) and Eq. (18), respectively.

exponentially rare, i.e., the probability for M particles to collide simultaneously decays like $(\sigma^3 n)^M$, with σ a characteristic interaction range. Hence, to extract the dominant thermodynamic effect of the interactions it suffices to calculate the radial distribution of one particle in the field created by a second particle, which may be located at the origin of the coordinate system, for a homogeneous system. Given the interpretation of $g(\mathbf{r})$ as a real space probability distribution, this problem is of course equivalent to the derivation of the barometer equation, with the known result (exercises) that $g(\mathbf{r})$ is just the Boltzmann factor[7]:

$$g(\mathbf{r}) \sim e^{-\beta\nu(\mathbf{r})} \qquad (n \to 0) \; . \tag{27}$$

Before a more formal discussion is provided, some immediate consequences of this shall be derived. Inserting Eq. (27) into the energy and virial equations of state, Eqs. (21), (26), one immediately deduces the equations of state of a dilute interacting gas[8]

$$\frac{U}{3Nk_BT/2} \sim 1 + \frac{n}{3} T\partial_T \int \mathrm{d}\mathbf{r}\, f(\mathbf{r}) = 1 - \frac{2n}{3} T\partial_T B(T) \tag{28}$$

and

$$\frac{p}{nk_BT} \sim 1 + \frac{n}{6} \int \mathrm{d}\mathbf{r}\, \mathbf{r}\cdot\nabla f(\mathbf{r}) = 1 + nB(T) \; , \tag{29}$$

respectively. The notation makes use of two important and common definitions, the *second virial coefficient*

$$B(T) \equiv -\frac{1}{2} \int \mathrm{d}\mathbf{r}\, f(\mathbf{r}) \; , \tag{30}$$

and the *Mayer function*

$$f(\mathbf{r}) \equiv e^{-\beta\nu(\mathbf{r})} - 1 \; , \tag{31}$$

which here plays the role of the function denoted by the same symbol in the related expansion at the end of the introductory statistical mechanics lecture. The Mayer function vanishes at large arguments for well-behaved interactions, thanks to the subtracted 1, which renders the above integrals non-extensive (which guarantees convergence in the thermodynamic limit). Also notice that the Mayer-function is the low-density limit of the pair correlation function, $h(\mathbf{r}) \sim f(\mathbf{r})$ for $n \to 0$. The second virial coefficient $B(T)$ is an effective volume that quantifies the leading corrections to the ideal gas limit of the constitutive equations. A positive/negative $B(T)$ (corresponding to effectively excluded/gained volume) is indicative of a predominantly repulsive/attractive interaction, which increases/decreases the pressure and the energy density, correspondingly.

---

[7] The symbol $\sim$ means "asymptotically equal", i.e. an exact equality in the indicated limit.

[8] For U, use $\beta\nu e^{-\beta\nu} = -\beta\partial_\beta f = T\partial_T f$ and the definitions of $f(\mathbf{r})$, $B(T)$; similarly for p.

A more formal derivation of the above results starts from the grand-canonical potential

$$\frac{pV}{k_BT} = \ln Z_G = \ln \sum_N z^N Z(N) \ , \tag{32}$$

where $Z(N) \equiv Z(V,T,N)$ is the canonical partition sum. Expanding Z_G in the fugacity $z \equiv e^{\beta\mu}$, which is a small quantity $z_{\mathrm{ig}} = n\lambda_T^3 \ll 1$ in the classical ideal-gas limit, one obtains a series with coefficients[9] that roughly increase like powers of $1/z$. The corresponding cumulant expansion (see exercises)

$$\ln Z_G = \ln[1 + Z(1)z + Z(2)z^2 \ldots] = Z(1)z + [Z(2) - Z(1)^2/2]z^2 + \ldots \tag{33}$$

is much better behaved. It not only has a small expansion parameter but also small expansion coefficients for a system close to the ideal gas limit. Note that the second order simply measures the deviation of the 2-particle partition sum $Z(2)$ from that of 2 independent particles, $Z(1)^2/2!$, and corresponding cancellations occur in all higher orders, as suggested by the observation that the leading term of the cumulant expansion already yields the exact result for an ideal gas after eliminating z in favor of $\langle N\rangle$,

$$\langle N\rangle = \partial_{\beta\mu} \ln Z_G = z\partial_z \ln Z_G = Z(1)z + [2Z(2) - Z(1)^2]z^2 + \ldots \tag{34}$$

For the leading terms in Eq. (32), i.e., neglecting $\mathcal{O}(Nn^2)$ in "…", one then has

$$\begin{aligned} \frac{pV}{k_BT} = \ln Z_G &= \langle N\rangle - [2Z(2) - Z(1)^2]\frac{z^2}{2} + \ldots \\ &= \langle N\rangle - \langle N\rangle^2 \frac{2Z(2) - Z(1)^2}{[Z(1)z + \ldots]^2}\frac{z^2}{2} + \ldots \end{aligned} \tag{35}$$

The series in powers of the particle concentration n thus obtained is called the *virial expansion*

$$\frac{p}{nk_BT} = 1 + B(T)n + \mathcal{O}(n^2) \tag{36}$$

with the second virial coefficient

$$B(T) = -\frac{V}{2}\left[\frac{Z(2)}{Z(1)^2/2} - 1\right] = -\frac{V}{2}[Q(2) - 1] = -\frac{1}{2}\int \mathrm{d}\mathbf{r}\, f(\mathbf{r}) \ , \tag{37}$$

where $Z_{\mathrm{ig}}(2) = Z^2(1)/2$ was used in the second step. This result demonstrates that it is indeed sufficient to solve a two-body problem, namely the two-particle configuration integral $Q(2)$, to compute the leading thermodynamic correction to the ideal gas. While the solution of a one-body problem was sufficient to describe a gas of N non-interacting particles, the solution of a two-body problem (which

---

[9] For the classical ideal gas $Z(1) = V/\lambda_T^3 = N/(n\lambda_T^3) = N/z_{\mathrm{ig}} \Rightarrow Z_{\mathrm{ig}}(N) = Z(1)^N/N! \approx z_{\mathrm{ig}}^{-N}$.

can always be reduced to a one-particle problem in an external field, *viz.* the barometer equation) suffices to deal with N rarely colliding particles.

In summary, the corrections to the thermodynamics of an ideal gas are for dilute systems completely contained in $B(T)$, which contains a two-body configuration integral. Assuming the general functional form of the interaction potential to be given — e.g. Lennard–Jones, (sticky) hard sphere, etc. — measurements of $B(T)$ can be used (and have excessively been used) to determine the potential parameters for atoms, simple molecules, colloidal particles or globular proteins. Dilute interacting gases are not more complex than ideal gases in an external field, i.e. the barometer equation. Calculations of higher order terms in the virial expansion require increasing effort and are of limited use, as the whole expansion will break down at the first really interesting occasion (e.g. a phase transition, or long-ranged interactions, which jeopardize the otherwise plausible fast convergence of the series).

Motivated by Eq. (27), one writes also the pair distribution $g(\mathbf{r})$ of a dense fluid in the form

$$g(\mathbf{r}) = e^{-\beta w(\mathbf{r})} \;, \tag{38}$$

which can be used to define the *potential of mean force* $w(\mathbf{r})$. With Eq. (14)

$$\nabla w(\mathbf{r}) = \frac{V}{\beta g(\mathbf{r})}\langle \nabla_{\mathbf{r}_{12}} \delta(\mathbf{r}-\mathbf{r}_{12})\rangle = \frac{V}{g(\mathbf{r})}\langle \delta(\mathbf{r}-\mathbf{r}_{12})\nabla_{\mathbf{r}_{12}}\nu(\mathbf{r}_{12})\rangle = \langle \nabla_{\mathbf{r}_{12}}\nu(\mathbf{r}_{12})\rangle_{\mathbf{r}_{12}=\mathbf{r}}$$

follows by partial integration, which explains the name and provides an alternative definition. From the last expression or — by analogy with the barometer formula — from Eq. (38), $w(\mathbf{r})$ is seen to be the reversible work needed to approach two infinitely distant particles of the fluid to a distance r through the "background solvent" provided by all the other particles. Splitting off the corresponding work $\nu(\mathbf{r})$ for the pair in empty space,

$$w(\mathbf{r}) = \nu(\mathbf{r}) + \Delta w(\mathbf{r}) \;, \tag{39}$$

one obtains the work — or, in fact, free energy — $\Delta w(\mathbf{r})$ isothermally absorbed by the solvent during this process, i.e. the part of the reversible work solely due to the presence of the solvent. In other words, the *other particles induce an effective interaction* $\Delta w(\mathbf{r})$ *between a chosen pair of test particles.* This is further discussed in the exercises for the example of the so-called "depletion attractions" induced in a fluid of purely repulsive particles.

The hard sphere fluid: paradigm of condensed matter

To illustrate the above formalism, it is useful to consider a fluid of hard spheres, for which

$$\nu(r) = \begin{cases} \infty & r<\sigma \\ 0 & r>\sigma \end{cases} \quad \Rightarrow \quad g(r) = \begin{cases} 0 & r<\sigma \\ e^{-\beta\Delta w(r)} & r>\sigma \end{cases} \tag{40}$$

This model idealizes the most salient feature of all atomic, molecular and most colloidal interactions, a strong hard-core repulsion at short center-of-mass distances, which plays the major role in determining the characteristic packing structure of most simple fluids. This is why the hard sphere fluid may justly be called a paradigm of condensed matter. Attractive interactions are often less pronounced than the hard-core repulsion and can be taken into account perturbatively once the hard sphere fluid is under control.

The only work in displacing a pair of hard spheres for distances $r > \sigma$ is due to the effective interactions $\Delta w(r)$ induced by the solvent, i.e. by the surrounding spheres. The function $\Delta w(r)$ can therefore be identified as the free energy cost for bringing two cavities of the size of the spheres from infinity to a distance r. The corresponding part of the radial distribution function $g_c(r) = e^{-\beta \Delta w(r)}$ is therefore also known as the "cavity distribution". Inserting Eq. (38) into the virial equation (26) yields a relation between the pressure and the contact value $g(\sigma)$ of $g(r)$,

$$\begin{aligned} \frac{p}{nk_BT} &= 1 - \frac{n}{6}\int d\mathbf{r}\, e^{-\beta\nu(r)-\beta\Delta w(r)}\beta\nu'(r)r \\ &= 1 + \frac{2\pi n}{3}\int dr\, \left(e^{-\beta\nu(r)}\right)' e^{-\beta\Delta w(r)} r^3 \\ &= 1 + \frac{2\pi n}{3}\int dr\, \delta(r-\sigma) g_c(r) r^3 \\ &= 1 + \frac{2\pi}{3} n\sigma^3 g(\sigma) = 1 + 4\phi g(\sigma)\ . \end{aligned} \tag{41}$$

Here, $\phi \equiv 4\pi/3 \cdot (\sigma/2)^3 n$ denotes the *volume fraction* of the spheres, and the final $g(\sigma)$ should be understood[10] as $g_c(\sigma) = g(r \to \sigma+)$. The take-home message is that the equation of state of a hard sphere fluid is determined by the contact value of the radial distribution function, i.e. by the number of collisions between the spheres. For concentrations $n \approx n_{\mathrm{cp}} = 1/v_{\mathrm{cp}}$ near close-packing[11], a simple free volume consideration suggests that the radial distribution function must develop a pointed next-neighbor peak of width $\Delta \approx (v - v_{\mathrm{cp}})/3\sigma^2$ at $r = \sigma$. The area under the peak,

$$4\pi n \int_\sigma^{\sigma+\Delta} dr\, r^2 g(r) \sim C \qquad (n \to n_{\mathrm{cp}}) \tag{42}$$

gives the number of nearest neighbors or "coordination number" C. For dense fluids and close-packed crystals $C \approx 12$. A simple estimate for the hard sphere equation of state $p(\phi)$ is therefore obtained by parametrizing the next-neighbor peak for $r \gtrsim \sigma$, e.g. by and exponential $g(\sigma)e^{-r/\Delta}$, plugging this into Eq. (42), and

---

[10] In contrast to g, g_c is smooth at $r = \sigma$ and no ambiguities due to discontinuity arise.

[11] In three dimensions, closest packing is attained in a fcc/hcp-crystal with $n = \sqrt{2}/\sigma^3$ corresponding to $\phi \approx 0.74$, as conjectured by J. Kepler, proved by K. F. Gauß with respect to all crystalline packings, and, at the end of the 20th century, by T. C. Hales in full generality.

solving for the contact value $g(\sigma)$ as a function of $n = 1/v$. More accurate semi-empirical formulas for the equation of state and the free energy of the hard-sphere fluid were proposed by Carnahan and Starling (and others) by extrapolating from the leading terms of the virial expansion:

$$\frac{p}{nk_BT} \approx 1 + \phi\frac{4-2\phi}{(1-\phi)^3}\,, \qquad \frac{F}{Nk_BT} \approx \frac{F_{\mathrm{ig}}}{Nk_BT} + \phi\frac{4-3\phi}{(1-\phi)^2}\,. \tag{43}$$

They can also be understood as results of a heuristic approach called "scaled particle theory" and, alternatively, the Percus–Yevick expression for $g(r)$, introduced further below. Reality is more complex. Already at $\phi \approx 0.58$ one observes a dramatic slowdown of the dynamics attributed to a glass transition that can not directly be guessed by simple free-volume arguments or any other of the mentioned theories, but is very well described by *mode-coupling theory*[12], a dynamic theory that builds on liquid-state theory. Moreover, there is a fluid-crystal phase coexistence with $p(n) = \text{constant}$ for $0.49 \lesssim \phi \lesssim 0.55$, which is best addressed by density functional theories, as outlined further below.

3 Attacking the many-body problem

3.1 Ornstein–Zernike integral equation

A central task in many-body theory is the prediction of phase transitions and the accompanying structural changes, as these provide not only the most spectacular applications but also the most significant and critical tests of the theory. The challenge then is to demonstrate how long-range correlations and divergent response coefficients emerge from a typically short-ranged interaction potential. At a critical point or at a spinodal, the pair correlation function $h(\mathbf{r}) = g(\mathbf{r}) - 1$ has to become long ranged such that its integral and, according to Eq. (18), the compressibility $S_{q\to 0}/(nk_BT)$ diverges within a narrow parameter range. A standard trick to guarantee that this subtle structure is preserved in practical calculations (which are, of course, approximate calculations) consists in introducing a new function $c(\mathbf{r})$ and its Fourier transform $c_{\mathbf{q}}$, called the *direct correlation function*, via

$$S_{\mathbf{q}} = 1 + nh_{\mathbf{q}} = \frac{1}{1 - nc_{\mathbf{q}}} \tag{44}$$

A small change in the density n then translates into singular behavior of $S_{q\to 0}$, and a finite error in $c_{\mathbf{q}}$ will only affect the precise location of the phase transition but not wreck its singular nature. This equation, and its transformation into real space

$$h(\mathbf{r}) = c(\mathbf{r}) + n\int d\mathbf{r}'\, c(\mathbf{r}-\mathbf{r}')h(\mathbf{r}')\,, \tag{45}$$

---

[12] W. Götze, Complex Dynamics of Glass-Forming Liquids: A Mode-Coupling Theory, Oxford Univ. Press, Oxford 2009.

are called *Ornstein–Zernike* (integral) equations (OZE). A formal justification of the pertinence of this structure and the interpretation of $c(\mathbf{r})$ are postponed to the next section, while the remainder of this section first attempts to provide an intuitive physical interpretation.

By expanding the denominator in Eq. (44) or iterating Eq. (45), respectively, one immediately sees that the OZE decomposes the correlations in $h(\mathbf{r})$ into chains of "direct correlations" as expressed by $c(\mathbf{r})$ (each $*$ stands for a convolution integral in real space or multiplication in Fourier space, respectively)

$$h = c + nc * c + n^2 c * c * c + \dots \; . \tag{46}$$

Indeed, the intuition of long-range correlations in $g(\mathbf{r})$ building up as chains of short-range correlations in $c(\mathbf{r})$ is an important guide in applications of the OZE. For long-ranged interactions, such as a Coulomb potential, the trick also works backwards and helps to explain how a long-ranged pair interaction gets screened by the presence of other charges. The chain structure of Eq. (46) suggests that the direct correlation function $c(\mathbf{r})$ should be a close relative of the pair potential $\nu(\mathbf{r})$, since it is the latter that helps to transmit interactions from particle to particle and thereby generates correlations between distant pairs of particles. However, in reality, these correlations are not merely transmitted by linear chains (there could be branches and loops in between). Therefore, the "dressed" interaction $c(\mathbf{r})$ and the "bare" interaction $\nu(\mathbf{r})$, although they play similar roles, are not identical.

There are essentially two different ways of applying Eq. (45) to a system of interest. First, as an exact self-consistency relation that has to be solved for the functions h and c, simultaneously. To this end, one has to supply closure conditions to constrain the two unknowns. Ideally, one would like to choose closures that avoid violations of any *a priori* known properties of $g(\mathbf{r})$, but there is no general systematic way how to find an exact closure. In practice, the self-consistent solutions are therefore non-perturbative, uncontrolled approximations corresponding to partial resummations of the virial series. A popular example is the Percus–Yevick closure for hard spheres,

$$h(r < \sigma) = -1 \qquad c(r > \sigma) = 0 \; . \tag{47}$$

While the first equation simply expresses the no-overlap condition, $g(r < \sigma) = 0$, and therefore is obviously exact, the second expresses the expectation that the direct correlation function should be of similar range as the interaction potential itself, i.e. it is a physically motivated approximation of *a priori* unknown quality. With the closure (and some effort), the OZE is analytically solvable and yields

$$c(r < \sigma) = \frac{1}{(1-\phi)^4}\left[6\phi(1+\phi/2)^2 r/\sigma - (1+2\phi)^2(1+\phi\, r^3/2\sigma^3)\right] , \tag{48}$$

which completes the information on $c(r)$ for an arbitrary volume fraction $\phi \equiv n\pi\sigma^3/6$ of the spheres. From this one gets pretty accurate approximations (see the yellow curve in Fig. 1) to the numerically exact $g(r)$ and S_q, which are known from Monte Carlo simulations (not shown). After integration, it also yields the Carnahan–Starling expression for the equation of state, Eq. (43).

The second reading of the OZE is as a kind of perturbation series that essentially "creates a highbrow $h(\mathbf{r})$ out of a lowbrow $c(\mathbf{r})$". This view can also provide some insight as to what might be a good closure. For example, one may start from the observation that, in the low-density limit,

$$n \to 0 \quad \Rightarrow \quad g(\mathbf{r}) - 1 \equiv h(\mathbf{r}) \sim c(\mathbf{r}) \sim f(\mathbf{r}) = e^{-\beta\nu(\mathbf{r})} - 1 \; . \tag{49}$$

Using this approximation for $c(\mathbf{r})$ in the OZE, one obtains the first correction to the limiting form for $h(\mathbf{r}) \sim f(\mathbf{r})$, namely

$$h(\mathbf{r}) \sim f(\mathbf{r}) + n \int \mathrm{d}\mathbf{r}' \, f(\mathbf{r} - \mathbf{r}') f(\mathbf{r}') \; . \tag{50}$$

In particular, for hard spheres $f(r > \sigma) = 0$ and $h(r > \sigma) = e^{-\beta\Delta w(r)} - 1 \sim -\beta\Delta w(r)$ is small at low densities, so that one immediately reads off

$$\beta\Delta w(\mathbf{r}_{12}) = -n \int \mathrm{d}\mathbf{r}_3 \, f(\mathbf{r}_{13}) f(\mathbf{r}_{23}) + \mathcal{O}(n^2) \; , \tag{51}$$

which amounts to an exponentiation of the right-hand side of Eq. (50). Inserting the hard-sphere Mayer function and doing the integral one thus obtains the interesting prediction that a pair of test spheres experiences an effective attraction ($\Delta w < 0$), called depletion attraction, due to the unbalanced pressure exerted by the surrounding particles (see exercises). The product of the two Mayer functions indicates that, to leading order, $\Delta w(\mathbf{r})$ takes the effect of a single additional particle onto the interaction of the pair of test particles into account. The leading order approximation therefore amounts to treating the "solvent" of other particles as an ideal gas, except for its interactions with the two test particles. This is known as the Asakura–Oosawa approximation in colloid science. A more accurate expression for Δw for hard-sphere suspensions that takes the interactions of the solvent particles into account, is easily reverse-engineered from Eqs. (38), (39) and the Percus–Yevick form for $g(r)$, of course.

To summarize, by an apparently simple rewriting (expressing h or S in terms of c), one has thus gained a robust and efficient approximation scheme for the otherwise forbiddingly complicated many-body problem encoded in a ridiculously high-dimensional nested configuration integral. The OZE involves only a single integration and the guessing of appropriate closures. With the OZE, already the simple limiting form of the direct correlation function for vanishing density (in fact, nothing but the good old barometer equation) yields interesting predictions for the pair correlations and the structure factor (see Fig. 1) that stay qualitatively trustworthy for finite (not too large) densities n and may even give a rough idea whereabout and how a phase transition might occur.

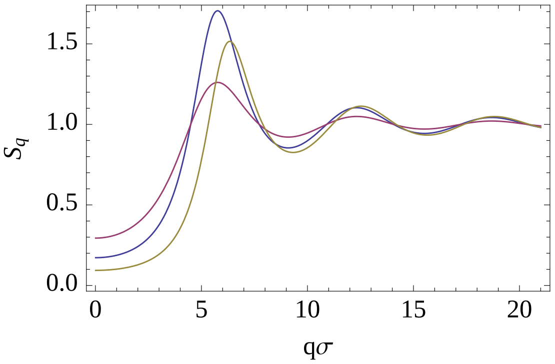

Figure 1: The structure factor of a hard sphere fluid of volume fraction $\phi = 0.3$ (purple) and $\phi = 0.6$ (blue) as estimated by the simplistic procedure described in the main text, which uses the hard-sphere Mayer function $f(r) = \theta(r - \sigma) - 1$ as an approximation for $c(\mathbf{r})$; see Eq. (50). While not quantitatively useful, it captures the qualitative features of the exact result (and, in fact, of the structure factor of any simple fluid) surprisingly well, as demonstrated by the comparison with the quite accurate Percus–Yevick result, Eq. (48) for $\phi = 0.3$ (yellow).

The following section establishes the above structure of pair correlations and direct correlations and their relation by the OZE on a more systematic basis, the double hierarchy of correlation functions generated from a twin couple of generating functionals generalizing the equilibrium free energy and grand-canonical potential.

3.2 Thermodynamic perturbation theory & FDT

The formalism to be introduced in this section provides the formal grounds for the Ornstein–Zernike equation and also for the discussion of the fluctuation-dissipation theorem and the density functional theories in the remainder of this part of the lecture.

The 2-lane stairway of correlations & response

This section generalizes the double hierarchy of isothermal derivatives of the equilibrium free energy $F(T, V, N)$ and grand-canonical potential $J(T, V, \mu)$, or rather their volume densities $\mathfrak{f}(T, n)$ and $\mathfrak{j}(T, \mu)$, respectively,

$$\begin{aligned} \partial_n \mathfrak{f} = \mu \ , \qquad & \partial_n^2 \beta \mathfrak{f} = \partial_n \beta\mu = (n^2 k_B T \kappa_T)^{-1} \\ -\partial_\mu \mathfrak{j} = n \ , \qquad & -\partial_{\beta\mu}^2 \beta \mathfrak{j} = \partial_{\beta\mu} n = n S_{q\to 0} \ , \end{aligned} \tag{52}$$

where β and V are understood to be constant throughout. The procedure can clearly be continued to higher order derivatives, which will however not be pursued in the following. Note that the compressibility sum rule, Eq. (18) is equivalent to the statement that the second derivatives of the mutually Legendre-transformed potentials are inverse to each other,

$$(\partial_n^2 \beta \mathfrak{f})(-\partial_{\beta\mu}^2 \beta \mathfrak{j}) = (\partial_n^2 \mathfrak{f})(-\partial_\mu^2 \mathfrak{j}) = 1\,. \tag{53}$$

This double hierarchy of derivatives of mutually Legendre transformed potentials becomes a very useful tool when it is upgraded to allow for spatial variations of all quantities involved except for the temperature $T = 1/k_B\beta$, which parametrizes the local equilibrium. The constant volume V is used to turn all extensive quantities into local densities. Spatial variations can arise from external forces that are represented by a potential $\mathcal{U}(\mathbf{r})$ or from the (pair-)interaction potential $\mathcal{V}(\mathbf{r})$ in the Hamiltonian. The latter can be rewritten in terms the microscopic densities,

$$\begin{aligned} H =& \sum_i \frac{\mathbf{p}_i^2}{2m} + \mathcal{V}(\{\mathbf{r}_i\}) + \mathcal{U}(\{\mathbf{r}_i\}) = \sum_i \frac{\mathbf{p}_i^2}{2m} + \sum_{i<j} \nu(\mathbf{r}_i - \mathbf{r}_j) + \sum_i u(\mathbf{r}_i) \\ =& \sum_i \frac{\mathbf{p}_i^2}{2m} + \frac{1}{2}\int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\, \hat{n}(\mathbf{r})\nu(\mathbf{r}-\mathbf{r}')\hat{n}(\mathbf{r}') - N\nu(0)/2 + \int \mathrm{d}\mathbf{r}\, \hat{n}(\mathbf{r})u(\mathbf{r}) \end{aligned} \tag{54}$$

The subtraction of $N\nu(0)$ from — and the prefactor 1/2 in front of — the interaction term correct for counting the self-interactions and for double counting all pairs, respectively. The aim is to write the free energies in a similar way, but with the microscopic densities $\hat{n}$ replaced by the ensemble averaged densities n (without hats). It turns out that a practical approximation for the free energy is indeed obtained by simply "forgetting the hats" in the Hamiltonian, which is indicative of the smooth evolution of the energy functional under coarse graining, but may initially cause some confusion.

The main difference of the local formalism thus obtained compared to the above formulas for homogeneous systems is that partial derivatives of the free energy densities turn into functional derivatives of the free energies themselves and that multiplications of free energies have to be interpreted as convolutions. (By Fourier transformation, one can always get rid of this complication). The benefit is that derivatives generate correlation and response functions and not merely moments of some global quantities and response coefficients, respectively.

To see how one arrives at such a formalism for the free energies, notice that the external potential has the form of a space-dependent generalization of the chemical potential term $-\mu N$ employed for homogeneous grand-canonical systems. The familiar structure is recovered for a constant (position-independent) potential $-\mu$. With the notion of a spatially heterogeneous generalized chemical potential (known as "electrochemical potential" in solid state physics)

$$\mu(\mathbf{r}) \equiv \mu - u(\mathbf{r}) \tag{55}$$

one can therefore delegate the thermodynamic consequences of the external potential to the grand-canonical potential in such a position-dependent chemical potential. For its functional derivatives with respect to $\mu(\mathbf{r})$ one finds by straightforward formal calculations using Eq. (54),

$$\frac{-\delta J}{\delta\mu(\mathbf{r})} = n(\mathbf{r}) \;, \quad \frac{-\delta^2 \beta J}{\delta\beta\mu(\mathbf{r})\delta\beta\mu(\mathbf{r}')} = \frac{\delta n(\mathbf{r})}{\delta\beta\mu(\mathbf{r}')} = \langle \delta\hat{n}(\mathbf{r})\delta\hat{n}(\mathbf{r}')\rangle \equiv G(\mathbf{r},\mathbf{r}') \;, \tag{56}$$

with $\delta\hat{n}(\mathbf{r}) \equiv \hat{n}(\mathbf{r}) - \langle\hat{n}(\mathbf{r})\rangle$. Also recall that the pair correlation function $G(\mathbf{r},\mathbf{r}')$ is essentially the Fourier transform of the structure factor (times particle number).

A corresponding "free-energy route" then follows by a Legendre transformation of the functional $J[\mu(\mathbf{r})]$ with respect to $\mu(\mathbf{r})$

$$F[n(\mathbf{r})] = J + \int \mathrm{d}\mathbf{r}\, n(\mathbf{r})\mu(\mathbf{r}) \;. \tag{57}$$

Since the external potential in the Hamiltonian was subsumed into the generalized chemical potential $\mu(\mathbf{r})$ in Eq. (55), this so-called *intrinsic free energy* belongs to the Hamiltonian without it. It clearly lacks the average potential energy

$$\langle \mathcal{U}(\{\mathbf{r}_i\})\rangle = \int \mathrm{d}\mathbf{r}\, \langle\hat{n}(\mathbf{r})\rangle u(\mathbf{r}) = \int \mathrm{d}\mathbf{r}\, n(\mathbf{r})u(\mathbf{r}) \tag{58}$$

contained in the Helmholtz free energy for Eq. (54) but subtracted in Eq. (57).

The functional derivatives of $F[n(\mathbf{r})]$ are[13]

$$\frac{\delta F}{\delta n(\mathbf{r})} = \mu(\mathbf{r}) \;, \quad \frac{\delta^2 \beta F}{\delta n(\mathbf{r})\delta n(\mathbf{r}')} = \frac{\delta\beta\mu(\mathbf{r})}{\delta n(\mathbf{r}')} = G^{-1}(\mathbf{r}',\mathbf{r}) \;, \tag{59}$$

so that

$$\int \mathrm{d}\mathbf{r}'' \, \frac{-\delta^2 J}{\delta\mu(\mathbf{r})\delta\mu(\mathbf{r}'')} \frac{\delta^2 F}{\delta n(\mathbf{r}')\delta n(\mathbf{r}'')} = \int \mathrm{d}\mathbf{r}''\, G(\mathbf{r},\mathbf{r}'')G^{-1}(\mathbf{r}'',\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}') \;. \tag{60}$$

It is often useful to isolate the exactly known ideal gas contribution or, more generally, any other "known" reference free energy. Up to gauge terms $\mathcal{O}(N)$ (such as the -1 inside the curly brackets), the free energy of the inhomogeneous ideal gas reads

$$\beta F_{\mathrm{ig}}[n(\mathbf{r})] = \int \mathrm{d}\mathbf{r}\, n(\mathbf{r})\{\ln[n(\mathbf{r})\lambda_T^3] - 1\} \tag{61}$$

with the derivatives

$$\frac{\delta F_{\mathrm{ig}}}{\delta n(\mathbf{r})} = k_B T \ln[n(\mathbf{r})\lambda_T^3] \equiv \mu_{\mathrm{ig}}(\mathbf{r}) \;, \qquad \frac{\delta^2 \beta F_{\mathrm{ig}}}{\delta n(\mathbf{r})\delta n(\mathbf{r}')} = \frac{1}{n(\mathbf{r})}\delta(\mathbf{r}-\mathbf{r}') \;. \tag{62}$$

---

[13] Do not worry about the distribution of $\mathbf{r}'$ and $\mathbf{r}''$ onto the first and second arguments of the functions involved. Second derivatives of free energies are interchangeable.

This suggests to introduce the so-called *excess free energy* $F_{\rm ex} \equiv F - F_{\rm ig}$, with the derivatives

$$\begin{aligned} \frac{\delta F_{\rm ex}}{\delta n(\mathbf{r})} &= \mu(\mathbf{r}) - \mu_{\rm ig}(\mathbf{r}) \equiv \mu_{\rm ex}(\mathbf{r}) \ , \\ \frac{\delta^2 \beta F_{\rm ex}}{\delta n(\mathbf{r})\delta n(\mathbf{r}')} &= \frac{\delta \beta\mu(\mathbf{r})}{\delta n(\mathbf{r}')} - \frac{1}{n(\mathbf{r})}\delta(\mathbf{r}-\mathbf{r}') = \frac{\delta\beta\mu_{\rm ex}(\mathbf{r})}{\delta n(\mathbf{r}')} \equiv -c(\mathbf{r},\mathbf{r}'). \end{aligned} \tag{63}$$

By comparison with the second line of Eq. (54), $c(\mathbf{r},\mathbf{r}')$ can be said to be the dressed version of the bare pair interaction $\nu(\mathbf{r}-\mathbf{r}')$ (the free energy looks like the Hamiltonian but with n and c in place of $\hat{n}$ and ν, respectively). The notation $c(\mathbf{r},\mathbf{r}')$ is not an accident. The quantity thus defined, or rather $n(\mathbf{r})c(\mathbf{r},\mathbf{r}')$, is seen to encode "excess" correlations beyond the δ-correlations of the ideal gas and reduces to the direct correlation function $n\,c(\mathbf{r}-\mathbf{r}')$, introduced above, in the special case of a translation invariant system. Inserting Eq. (63) together with the (total) *pair correlation function*[14]

$$n(\mathbf{r})n(\mathbf{r}')h(\mathbf{r},\mathbf{r}') \equiv n(\mathbf{r},\mathbf{r}') - n(\mathbf{r})n(\mathbf{r}') = G(\mathbf{r},\mathbf{r}') - n(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}') \ , \tag{64}$$

into Eq.(60), one finds the familiar structure (exercises):

$$h(\mathbf{r},\mathbf{r}') = c(\mathbf{r},\mathbf{r}') + \int \mathrm{d}\mathbf{r}''\, n(\mathbf{r}'')c(\mathbf{r}'',\mathbf{r})h(\mathbf{r}',\mathbf{r}'') \ . \tag{65}$$

This is, of course, nothing but the Ornstein–Zernike equation in new garment. The earlier versions are immediately recovered for a homogeneous system with constant density $n(\mathbf{r}) = n$, for which also the definition of $h(\mathbf{r},\mathbf{r}')$ reduces to that of $h(\mathbf{r}-\mathbf{r}')$.

This now further clarifies the pertinence of the OZE and the notion of the direct correlation function, and shows that both are deeply ingrained in the general structure of equilibrium thermodynamics and statistical mechanics, providing a paradigm for quantum field theories and non-equilibrium statistical mechanics. The OZE constitutes a relation between the 2-point correlation and/or response functions of different hierarchies, one obtained along the free-energy route, the other along the grand-potential route. In terms of response functions, it simply states the (mathematically trivial) reciprocity of $\partial_\mu n$ and $\partial_n \mu$ in terms of the excess fluctuations and response. The response of the density to an infinitesimal fluctuation of the chemical potential is the reciprocal of the response of the chemical potential to an infinitesimal density fluctuation. As the above discussion tried to convey, despite this seemingly trivial content, the OZE can serve as a powerful passkey to the otherwise forbiddingly complicated many-body problem.

---

[14] Here, $n(\mathbf{r},\mathbf{r}')$ is the usual 2-point correlation function without self-correlations.

Fluctuation-response theorem

As already pointed out, the second derivatives of the generalized thermodynamic functionals $F[n(\mathbf{r})]$ and $J[\mu(\mathbf{r})]$ in Eqs. (56) and (59) can be interpreted as spatially varying susceptibility functions or correlation functions, respectively. These equations thus generalize results about the equivalence of fluctuations and response coefficients; e.g. the relation between grand-canonical number fluctuations and the compressibility derived in introductory texts on statistical mechanics, or the compressibility sum rule in Eq. (18). They can also be understood as a local, differential version of the virial equation from Sec. 2.1 for spatially variable pressure. The relation between packing structure, thermal fluctuations, and macroscopic material properties is thereby cast into a form that includes detailed spatial information on both the structural and material side, to leading order. To bring it into a more familiar form, which also exposes more clearly its perturbative nature, it shall now be re-derived for the special case of a homogeneous fluid. For the sake of the argument, the external perturbation is assumed to be periodic (a Fourier component of a weak but otherwise arbitrary perturbing electric or gravitational potential, say),

$$u(\mathbf{r}) = \frac{u_q}{V} e^{i\mathbf{q}\cdot\mathbf{r}} \; . \tag{66}$$

This corresponds to the perturbation Hamiltonian $\mathcal{U} = u_q \hat{n}_{-q}/V$, with $\hat{n}_q = \sum_i e^{-i\mathbf{q}\cdot\mathbf{r}_i}$ the Fourier transformed microscopic density. Then, the resulting density field $\langle \hat{n}_q \rangle_{\mathcal{U}}$ in presence of the perturbation $\mathcal{U}$ is obtained to leading order in the perturbing field as

$$\langle \hat{n}_q \rangle_{\mathcal{U}} = \frac{\langle \hat{n}_q e^{-\beta\mathcal{U}} \rangle}{\langle e^{-\beta\mathcal{U}} \rangle} \sim \frac{\langle \hat{n}_q (1-\beta\mathcal{U}) \rangle}{\langle 1-\beta\mathcal{U} \rangle} \sim n_q + \frac{\beta u_q}{V} \Big(\langle \hat{n}_q \rangle \langle \hat{n}_{-q} \rangle - \langle \hat{n}_q \hat{n}_{-q} \rangle \Big) \; .$$

Here and in the following, quantities calculated in, or pertaining to the perturbed state are marked with $\mathcal{U}$ to distinguish them from corresponding quantities in the reference fluid. The strength of the deviations from the homogeneous reference state are controlled by the structure factor S_q of the unperturbed system, which—at the same time—characterizes the fluctuations and packing structure,[15]

$$\Delta n_q^{(\mathcal{U})} \equiv \langle \hat{n}_q \rangle_{\mathcal{U}} - \langle \hat{n}_q \rangle = -\frac{\beta u_q}{V} \langle \delta\hat{n}_q \delta\hat{n}_{-q} \rangle = -\beta u_q n S_q \; . \tag{67}$$

This is a typical *linear-response* result. The response of the density is linearly proportional to the perturbing field u_q. The prefactor is an unperturbed equilibrium average independent of u_q, hence characteristic of the equilibrium system. The local susceptibility χ generalizing the overall isothermal compressibility κ_T

---

[15] Note the overloading of the symbols S_q ($q = 0$ contribution included/excluded) and δ (Dirac delta, functional derivative and variation $\delta\hat{n} \equiv \hat{n} - n$ around a local/global average $n \equiv \langle \hat{n} \rangle$).

is the functional derivative of the local density with respect to the perturbing potential. For a homogeneous system, it turns into an ordinary partial derivative in Fourier space,

$$\chi(\mathbf{r},\mathbf{r}') \equiv \delta n(\mathbf{r})/\delta u(\mathbf{r}') = -\beta G(\mathbf{r},\mathbf{r}') \;, \qquad \chi_q = \partial n_q/\partial u_q = -\beta n S_q \;. \tag{68}$$

The structure factor, the two-point correlation function of the density fluctuations, is at the same time the susceptibility controlling the density change in response to an external perturbing potential. This is the *fluctuation-response theorem*, which contains the compressibility sum rule as the special case $q \to 0$, and which is itself a specialization (to temporally stationary situations) of the more general (dynamic) fluctuation-dissipation theorem (FDT).

Perturbation theory

Perturbation theories can be used to systematically improve a "known" reference system, such as the ideal gas or the hard sphere fluid. The ultimate aim would be to calculate the excess free energy. The fluctuation-response relation derived in the preceding paragraph concerns the linear response to (or first-order perturbation by) an external potential. This can be integrated to obtain a "perturbed" free energy that contains corrections to the reference system. Similarly, starting from the first Eq. (63) (at vanishing density), one can integrate[16] over the strength α of a given fixed density profile $n_\alpha(\mathbf{r}) \equiv \alpha n(\mathbf{r})$, namely

$$\beta F_{\mathrm{ex}} = \int_0^1 \mathrm{d}\alpha \int \mathrm{d}\mathbf{r}\; n(\mathbf{r}) \mu_\alpha^{\mathrm{ex}}(\mathbf{r}) \;. \tag{69}$$

The index α refers to the implicit dependence of $\mu_\alpha^{\mathrm{ex}}(\mathbf{r})$ on the density $n_\alpha = \alpha n$ and keeps track of the accumulating changes in the packing structure (reflected in μ^{ex}) upon increasing it. If you loosely think of the "charging parameter" α as the amount of air pumped into a dinghy of shape $n(\mathbf{r})$ to inflate it, the index α of the generalized chemical potential $\mu_\alpha^{\mathrm{ex}}(\mathbf{r})$ reflects the increase of the resisting elastic force in the rubber upon stretching, so to say.

Starting from the second Eq. (63), instead, one has to integrate the direct correlation function from the reference state to the state with the desired density n of excitations. Since there are now two factors of the density, one has to integrate twice:

$$\beta F_{\mathrm{ex}}[n(\mathbf{r})] = -\int_0^1 \mathrm{d}\alpha \int \mathrm{d}\mathbf{r}\, n(\mathbf{r}) \int_0^\alpha \mathrm{d}\alpha' \int \mathrm{d}\mathbf{r}'\, n(\mathbf{r}')\, c_{\alpha'}(\mathbf{r},\mathbf{r}') \;. \tag{70}$$

The first integration produces a non-trivial force at a finite density $n\alpha$ that has then to be integrated up over the differential response $n\mathrm{d}\alpha$ to get the free energy.

---

[16] Do not confuse this with functional integration over the functional argument of $F_{\mathrm{ex}}[n(\mathbf{r})]$.

The dependence of the direct correlation function on α again tracks the accumulating changes in the pair structure upon increasing αn from zero to n. For a homogeneous final state, using the identity

$$\int_0^1 \mathrm{d}\alpha \int_0^\alpha \mathrm{d}\alpha'\, c_{\alpha'} = \int_0^1 \mathrm{d}\alpha\,(1-\alpha)c_\alpha \;, \tag{71}$$

this simplifies to

$$\beta\mathfrak{f}_{\mathrm{ex}} = n^2 \int_0^1 \mathrm{d}\alpha\,(\alpha-1) \int \mathrm{d}\mathbf{r}\, c_\alpha(\mathbf{r}) \;. \tag{72}$$

A similar game can be played by blowing up the interaction potential $\nu(\mathbf{r},\mathbf{r}')$ rather than the external potential or the density. Starting from

$$\frac{\delta J}{\delta\nu(\mathbf{r}-\mathbf{r}')} = \frac{\delta F}{\delta\nu(\mathbf{r}-\mathbf{r}')} = \frac{1}{2} n(\mathbf{r},\mathbf{r}') \;, \tag{73}$$

if the interaction $\nu_\alpha = \nu_0 + \alpha(\nu-\nu_0)$ (with $\nu_0 = 0$ in the simplest case) is switched on by a charging parameter α, one has

$$F_{\mathrm{ex}}[n(\mathbf{r})] = \frac{1}{2} \int_0^1 \mathrm{d}\alpha \int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\, n_\alpha(\mathbf{r},\mathbf{r}')\nu_\alpha(\mathbf{r},\mathbf{r}') \;. \tag{74}$$

The problem with the exact relations Eqs. (70), (74) is that c_α and n_α are of course not known, *a priori*. As already mentioned in the context of the constitutive equations, one needs expressions for the radial distribution in a whole continuous parameter range to infer the thermodynamics from the packing structure. In some practical situations, one might be satisfied with a less ambitious program, though. First order perturbation results are for instance obtained by replacing the unknown pair functions in Eqs. (70), (74) by those pertaining to the reference state. This is reminiscent of the first order corrections to energy levels in quantum mechanics, which are given by the matrix elements of the perturbation Hamiltonian in the basis functions of the unperturbed problem.

Further, the so-called RPA ("random phase approximation") replaces Eqs. (70), (74) by[17]

$$F_{\mathrm{ex}}[n(\mathbf{r})] = \frac{1}{2} \int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\, n(\mathbf{r})\nu(\mathbf{r}-\mathbf{r}')n(\mathbf{r}') \qquad \text{(RPA)} \;. \tag{75}$$

This ubiquitous approximation can be interpreted as arising from dropping the $\alpha-$integrations and either of the following two simplifications

$$n(\mathbf{r},\mathbf{r}') \to n(\mathbf{r})n(\mathbf{r}') \qquad \text{or} \qquad c(\mathbf{r},\mathbf{r}') \to -\beta\nu(\mathbf{r},\mathbf{r}') \;. \tag{76}$$

The first is known as the decoupling approximation, which shows that the RPA becomes exact at large distances in a (non-critical) fluid, where the correlations

---

[17] The homogeneous contribution $n^2\nu_{q\to 0}$ to the integrand is sometimes excluded by replacing the density fields $n(\mathbf{r})$ by their deviations $\delta n(\mathbf{r})$ from the homogeneous reference state $n(\mathbf{r}) = n$.

decay over a finite length scale. The second amounts to replacing the dressed by the bare interaction. It has already been used above in a slightly upgraded version — namely in the form $c(\mathbf{r}) \to f(\mathbf{r}) = e^{-\beta\nu(\mathbf{r})} - 1$, which amounts to a different partial resummation of the virial series (see Fig. 1). There, it was motivated as a low-density approximation. From the first interpretation in Eq. (76), it is now seen to also hold for dense fluids at large distances $|\mathbf{r} - \mathbf{r}'| \to \infty$. In particular, Eq. (76) corroborates that $c(r)$ and $\nu(r)$ are of similar range in a fluid, where $n(\mathbf{r}, \mathbf{r}') \to n(\mathbf{r})n(\mathbf{r}')$ for $|\mathbf{r} - \mathbf{r}'| \gg \sigma$. This property is an important ingredient in all successful closure schemes for the OZE of fluids.

Finally, the thermodynamic minimum property of the grand potential (the 2^{nd} law),

$$\frac{\delta J}{\delta n(\mathbf{r})} = \frac{\delta F}{\delta n(\mathbf{r})} - \mu(\mathbf{r}) = \frac{\delta F_{\text{ex}}}{\delta n(\mathbf{r})} - \mu_{\text{ex}}(\mathbf{r}) = 0 \ , \qquad (77)$$

can be exploited to venture beyond the above approximations by a variational principle. Free energies calculated with approximate structure functions (or based on guessed trial functions) can then *a posteriori* be minimized with respect to free parameters. Similarly, the independence of the free energies of artificial decompositions of the interactions into a reference part ν_0 and perturbation part $\nu - \nu_0$ can be exploited. All of the above techniques (and many others) are used in so-called density functional theories.

A schematic sketch of how these theories work in principle (not literally) is provided in Fig. 2, think of a fluid of interacting particles, where the interactions can be split into an effective hard-sphere part and a perturbation part. The hard-sphere part is dealt with in the Percus–Yevick approximation to obtain $c(r)$ analytically via the Ornstein-Zernike equation, while the perturbation part is dealt with in perturbation theory, e.g. in the RPA approximation. Finally, the free energy obtained by integration is minimized to find an optimum value for some artificial model parameters (schematically represented by the effective hard-core diameter σ in the figure). To put such ideas to work in practice is an intricate story of its own. The next section only provides a very brief outline.

3.3 Density-functional theories

Density-functional theory (DFT) is a very successful and popular framework based on the above formalism that extends (and includes as a special case) the well-known Ginzburg–Landau theory discussed in introductory texts, but also theories of freezing and melting transitions (in simple fluids as well as in liquid crystals), and the Poisson–Boltzmann and Debye–Hückel theory. Appropriately extended versions of DFT involving particle exchange effects are very useful in quantum mechanical many-body problems, e.g. in describing electrons and quasi-particles in (semi-)conductors and large atoms. In the following, the emphasis is on a few paradigmatic examples of fundamental interest in many areas of physics.

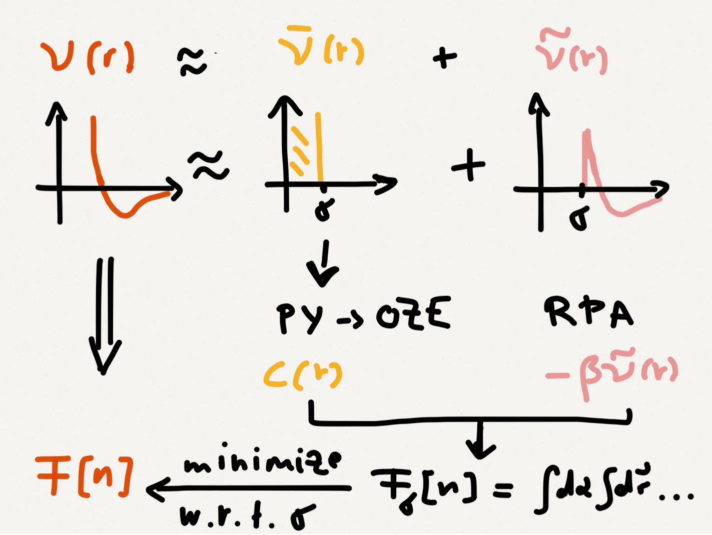

Figure 2: Schematic of how the various perturbative and non-perturbative approximation methods introduced so far — hard-sphere-fluid, PY, OZE, thermodynamic integration, RPA, free energy minimization — can (in principle) be combined to produce approximate expressions for the free energy functional $F[n]$ of a dense interacting fluid with pair interaction potential $\nu(r)$.

Inhomogeneous ideal gas

Basically the only exactly known density functional of an inhomogeneous fluid is that of the inhomogeneous ideal gas. It is obtained by replacing the constant density n by the spatially varying field $n(\mathbf{r})$ in the known bulk formula,

$$\beta F_{\mathrm{ig}}[n(\mathbf{r})] = \int \mathrm{d}\mathbf{r}\, n(\mathbf{r}) \big(\ln[n(\mathbf{r})\lambda_T^3] - 1\big) \;\Rightarrow\; \frac{\beta\delta^2 F_{\mathrm{ig}}}{\delta n(\mathbf{r})\delta n(\mathbf{r}')} = \frac{1}{n(\mathbf{r})}\delta(\mathbf{r}-\mathbf{r}')\,. \qquad (78)$$

In the homogeneous state the fields degenerate to $n(\mathbf{r}) = n$, $G(\mathbf{r},\mathbf{r}') = n\delta(\mathbf{r}-\mathbf{r}')$, and hence $ng(\mathbf{r}) = n$, $h(\mathbf{r}) = c(\mathbf{r}) = 0$, and $S_q \equiv 1$.

Slightly inhomogeneous fluids

Even for a strongly interacting many-body system, weak fluctuations around a homogeneous reference state may often be treated as a weakly interacting gas. The corresponding model of a slightly inhomogeneous fluid is among the most commonly applied models in many-body physics. It is the (classical) basis for the quantized elementary excitations and quasi-particles that govern the world of solid-state physics, but it is equally useful in soft matter physics. The inhomogeneous part of the free energy in terms of the weak density fluctuations

$\delta n(\mathbf{r}) = n(\mathbf{r}) - n$ with $|\delta n(\mathbf{r})| \ll n$ reads[18]

$$\beta F_{\text{ih}} = \int \mathrm{d}\mathbf{r}\ \{n(\mathbf{r}) \ln[n(\mathbf{r})/n] - \delta n(\mathbf{r})\} - \frac{1}{2} \int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\, c(\mathbf{r}, \mathbf{r}')\delta n(\mathbf{r})\delta n(\mathbf{r}')\ . \tag{79}$$

Note that due to the assumed weakness of the fluctuations, there is no additional integration over a "charging parameter" α involved, here. This is simply a rewriting of the second line of Eq. (63) as a differential form for infinitesimal $\delta n(\mathbf{r})$. The direct correlation function of the reference state is assumed to be known from somewhere (e.g. from the OZE) or estimated via the RPA. A familiar example for such theories is the square-gradient expansion familiar from the Ginzburg–Landau free energy for critical fluctuations, that amounts to the additional small wave vector approximation $c_q \sim c_0 + c_1 q^2/2$. (Here, $c_1/4\beta$ would become the coefficient in front of the square gradient term in the free energy after Fourier back transformation.) But many other remarkably successful DFTs — in fact all those introduced here — are constructed in the same way.

For the inhomogeneous free energy, Eq. (79), the minimum condition, Eq. (77), reads

$$\frac{\beta \delta F_{\text{ih}}}{\delta n(\mathbf{r})} = \beta\mu_{\text{ih}}(\mathbf{r}) = \ln[n(\mathbf{r})/n] - \int \mathrm{d}\mathbf{r}'\, c(\mathbf{r}, \mathbf{r}')\delta n(\mathbf{r}')\ . \tag{80}$$

This has to vanish or to balance an external potential $-\beta u(\mathbf{r})$, if present. (The overall chemical potential μ serves to fix the value of n.) In other words, the spatial density profile $n(\mathbf{r})$ obeys the self-consistency relation

$$n(\mathbf{r}) = n e^{-\beta u_{\text{eff}}(\mathbf{r})} \quad \text{with} \quad \beta u_{\text{eff}}(\mathbf{r}) = \beta u(\mathbf{r}) - \int \mathrm{d}\mathbf{r}'\, c(\mathbf{r}, \mathbf{r}')\delta n(\mathbf{r}') \tag{81}$$

a position dependent effective potential, which is self-generated by the interactions. This generalizes the well-known self-consistency equation, most often first encountered in the mean-field theory for the Ising model in introductory statistical-mechanics texts, to a spatially varying mean field. It incidentally improves it by replacing the bare pair interaction potential with the direct correlation function. (In return, one learns that, if mean-field theory is exact in infinite dimensions and/or for van-der-Waals-type pair interactions that couple all possible pairs with equal strength, then the same must hold for the RPA, which further supports the robust character of this approximation.) For vanishing interactions the barometer equation is recovered. The calculation of weak density fluctuations in an interacting fluid thus parallels the solution of the barometer equation for an ideal gas with the additional (hard) problem of finding the self-consistent self-generated mean field (the last term in the equation). If the direct correlation function is not yet known (e.g. from the OZE), it is commonly replaced by the bare potential (or by the Mayer function), in the RPA approximation. Even then,

---

[18] The last term in the curly brackets is a gauge term (actually vanishing upon integration) kept to make later cancellations with the first term explicit.

the Eq. (81) represents an infinite-dimensional nonlinear eigenvalue problem that usually cannot be solved exactly by analytical means.

Example: Poisson–Boltzmann theory

A widely known and applied example for the above formalism is the Poisson–Boltzmann theory (also known by the name Guoy–Chapman theory to chemists). It corresponds to the RPA approximation of the self-consistency equation (81), namely the replacement of the direct correlation function by the bare Coulomb potential, $-c(r) \to \beta\nu(r) = z^2\ell_B/r$. The effective potential βu_{eff} is then simply the electrostatic potential (as felt by a charge $\pm ze$)

$$\beta u_{\text{eff}} \equiv \pm\beta\phi(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \, \frac{\pm z^2\ell_B}{|\mathbf{r}-\mathbf{r}'|} n_e(\mathbf{r}') \;, \tag{82}$$

arising from the charge fluctuations $n_e(\mathbf{r}) \equiv n_+(\mathbf{r}) - n_-(\mathbf{r})$ in the globally neutral plasma of particles of charge $\pm ze$ and local densities $n_\pm(\mathbf{r})$. The characteristic length scale $\ell_B \equiv \beta e^2$, called *Bjerrum length*, is defined by the balance of electrostatic and thermal energy for two elementary charges at distance ℓ_B (i.e., $k_BT = e^2/\ell_B$).

The name Poisson–Boltzmann theory derives from the following obvious interpretation. The electrostatic potential $\phi(\mathbf{r})$ generated by the charge density $n_e(\mathbf{r})$ is a solution of Poisson's equation

$$\nabla^2\beta\phi(\mathbf{r}) = -4\pi z^2\ell_B n_e(\mathbf{r}) = -\varkappa^2[n_+(\mathbf{r}) - n_-(\mathbf{r})]/n \qquad \text{(Poisson)} \;. \tag{83}$$

Here, the *Debye screening length* $\varkappa^{-1}$, with $\varkappa^2 \equiv 4\pi z^2\ell_B n$, appears as a natural scale. From the Poisson equation, the Poisson–Boltzmann equation

$$\nabla^2\beta\phi(\mathbf{r}) = \varkappa^2 \sinh\beta\phi(\mathbf{r}) \qquad \text{(Poisson–Boltzmann)} \tag{84}$$

follows by simply estimating the local densities of the positive and negative charges on the right from the barometer formula (the Boltzmann factors) for the self-generated potential $\phi(\mathbf{r})$, namely,

$$n_\pm(\mathbf{r}) = (n/2)e^{\mp\beta\phi(\mathbf{r})} \qquad \text{(barometer)} \;. \tag{85}$$

The physical significance of the Debye length is clearly revealed if one linearizes Eq. (84). In this so-called Debye–Hückel approximation ($\beta\phi \ll 1$) it is seen to be the distance over which a charge mismatch is compensated by the polarization of the surrounding plasma. As a consequence, the long-ranged bare Coulomb interaction is screened beyond distances $r \approx \varkappa^{-1}$ by charge fluctuations, and the "dressed" interactions decay exponentially $\phi(r) \propto e^{-\varkappa r}/r$ rather than reciprocally with the distance r (in 3 dimensional space).

The above derivation of the Poisson–Boltzmann equation clearly exposes the mean-field character of the theory, as it treats the particles as an ideal gas in the self-generated mean field $\phi(\mathbf{r})$, without considering more subtle inter-particle correlations, which is bound to fail for a not so weakly inhomogeneous system.

The RPA charge structure factor and fluctuation function are readily obtained directly from the OZE expression for the structure factor, namely

$$S^e_q \overset{\text{RPA}}{\approx} \frac{1}{1+n\beta\nu_q} = \frac{q^2}{\varkappa^2+q^2} \qquad \Longrightarrow \qquad -h^e_q = 1 - S^e_q = \frac{\varkappa^2}{\varkappa^2+q^2} \tag{86}$$

The vanishing of the "charge compressibility" $S^e_{q\to 0}$ is clearly equivalent to global charge neutrality. Fourier back transformation Eq. (86) yields the Debye–Hückel form of $h^e(r)$, which can, to the present order of approximation, be identified with the (charge) potential of mean force $\beta w^e(r)$

$$-h^e(r) \approx \beta w^e(r) \approx \frac{z^2\ell_b}{r}\exp(-\varkappa r)\ . \tag{87}$$

So the actual interaction potential between two charge fluctuations embedded in a plasma is not the bare Coulomb potential, but a screened (or "dressed") version. It is Coulombic only at short distances $r \ll \varkappa^{-1}$ but vanishes at large distances $r \gg \varkappa^{-1}$. A charge fluctuation is thus not visible from large distances. The result can also be interpreted in terms of the "solvent contributions" encoded in $\Delta w(\mathbf{r}) = z^2\ell_b(e^{-\varkappa r}-1)/r$, which suppress the tail of the bare Coulomb potential by creating a polarization cloud or correlation hole around any net charge. The screening by this cloud becomes more effective with higher plasma density n, but is somewhat "washed out" at high temperature T, which is reflected in the $T-$ and $n-$dependencies of $\varkappa$.

Freezing and mesophases

Other significant early examples for DFTs were the Thomas–Fermi theory for atoms with many electrons and Onsager's theory of a nematic transition in suspensions of hard rods. Onsager's theory[19] from 1949 is particularly interesting, since it combines two independent theoretical concepts introduced above, the virial expansion and density-functional theory. It is a paradigm of partial ordering (orientational ordering without positional ordering) into a so-called micro- or mesophase, a theme with countless variations in (soft) condensed matter theory; and finally it is an early example for a purely entropic or "geometric" transition. Thereby, it anticipates later theories of freezing and melting of simple fluids.

The essential physics of the nematic transition is that isotropic rods of length L and diameter a jam up at high densities, as familiar from the mikado game. To avoid an entropy crisis, i.e., a complete loss of conformational entropy at the jamming density[20] $n = n_c \approx 1/L^2a$, the system sacrifices a bit of its orientational

---

[19] L. Onsager, Annals of the New York Academy of Sciences 51 (627-659) 1949.

[20] Onsager argues that, by virtue of the low critical volume fraction $a^2Ln_c \simeq a/L$, the virial series can be truncated after the second virial coefficient and the direct correlation function can be approximated by the Mayer function, both with negligible error, for very thin needles.

disorder to retain some more of its translational freedom. This trick only works if a macroscopic fraction of the molecules can agree on an average preferred orientation: spontaneous long-range orientational order has to emerge in order that the center-of-mass coordinates can retain their fluid-like distribution (whence the name *liquid crystal*, referring to the lack of positional long-range order).

In 1979, Ramakrishnan und Yussouff used Eq. (79) along similar lines with good success to elucidate the geometric nature of the freezing and melting transitions in classical fluids, explain the universality of the packing structure near these transitions, and numerically compute quantitative predictions for the phase diagram. Their analysis provided a theoretical basis for successful phenomenological rules of thumb, such as the Verlet rule and the Lindemann criterion, which are widely used in practical applications. The *Lindemann criterion* states that crystals melt if the mean-square fluctuations of the atoms exceed 10% of their average next-neighbor distance (the lattice unit), and the *Hansen–Verlet rule* states that fluids freeze if the height of the main peak in their structure factor exceeds 3, thereby hinting at different underlying mechanisms for both transitions.

The DFT approach regards the crystal as a (not so) slightly inhomogeneous fluid. The procedure can be characterized as "fishing for crystal structures". Briefly, one looks for instances of a non-zero order parameter $n_{\mathbf{K}_i} \equiv \langle \sum_j e^{i\mathbf{K}_i \cdot \mathbf{r}_j} \rangle = \mathcal{O}(N)$ (i.e. Bragg peaks) in a preferably exhaustive set of possible reciprocal lattice vectors $\mathbf{K}_i$. To this end one needs an ansatz. Typically, a sum of Gaussian distributions of width a centered at the lattice sites $\{\mathbf{R}\}$

$$n(\mathbf{r}) = n_s \big(\sqrt{\pi} a / a_0\big)^{-3} \sum_{\{\mathbf{R}\}} e^{-(\mathbf{r}-\mathbf{R})^2/2a^2} \tag{88}$$

is plugged into Eq. (79) as a trial function. Local minima of the resulting free energy other than the homogeneous fluid reference state are then searched by varying the model parameters n_s, a_0, and a for a prescribed trial crystallographic lattice $\{\mathbf{R}(a_0)\}$ such as BBC, FCC, etc., and the one with the lowest free energy is then proposed as the (most likely) equilibrium structure. While the lattice constant a_0 refers to the abstract lattice, the density n_s refers to the actual mass sitting on this lattice. If it is treated as an independent parameter (not tied to $1/a_0^3$, say), the ansatz can account even for vacancies.

Limitations of DFT

The above examples illustrate that DFT can be a quick and easy way to get impressive results by comparably little effort — essentially by dropping the magic hats on the densities $\hat{n}(\mathbf{r})$ in the Hamiltonian to accomplish much of the impossible task of calculating the partition sum of an interacting many-body system. The trick works remarkably well, even for some pretty complex first-order phase transitions. Are we in paradise, then, or what are the limitations of this strikingly simple and powerful recipe? Even though DFT exploits and builds on formally exact relations and admits fluctuations via the spatially varying density field $n(\mathbf{r})$, it is mean-field like (as van-der-Waals or Curie-Weiß theory), in practice. The approximations required to address real-life applications (such as weak inhomogeneity, RPA) usually introduce uncontrolled errors that spoil the critical behavior at continuous phase transitions and other subtle correlations. Briefly, from the structure of the OZE and the compressibility sum rule

$$n\kappa_T = \beta S_{q\to 0} = \frac{\beta}{1 - nc_{q\to 0}} \overset{\text{RPA}}{\approx} \frac{\beta}{1 + n\beta\nu_{q\to 0}} = \frac{k_B^{-1}}{T - T_c}\ . \tag{89}$$

So the compressibility κ_T can clearly be made to diverge like $(T - T_c)^{-\gamma}$ with $\gamma = 1$ at a critical temperature $T_c \equiv -n\nu_{q\to 0}/k_B$. But it remains obscure how γ might ever acquire a non-mean-field (non-rational) value — as phenomenologically observed — without extraordinary conspiracy. This discrepancy turned into an outstanding historical challenge for theoretical physics that was resolved in the 1970's, as discussed in the next section. Another limitation of DFT is that it usually does not get rid of some guessing of trial functions or, in more elaborate versions, at least of some free parameters. With respect to the task of obtaining numerically precise predictions this freedom may suit the practitioner, while it is somewhat less satisfying with respect to the task of generating insight into fundamental physical mechanisms.

4 Critical phenomena and renormalization

4.1 Critical behavior and Landau theory

Critical exponents

Critical behavior is the thermodynamic behavior associated with continuous phase transitions, as typically found at an end of a discontinuous phase transition lines (such as the liquid-vapor coexistence line in a $p - T-$diagram). Generic example are provided by the liquid-vapor critical point or by a periodic spin lattice that responds paramagnetically at high temperature but turns into a ferromagnet below the Curie temperature. The Bose–Einstein condensation from the introductory lecture is a somewhat trivial example, since the model is interaction-free. Quite generally, one characterizes critical behavior phenomenologically by the exponents of the power-law singularities of various thermodynamic quantities of interest at vanishing critical parameter $t \equiv T/T_c - 1$ (or $t \equiv 1 - T_c/T$),

$$C_V \propto |t|^{-\alpha} \tag{90}$$

$$\psi \propto |t|^{\beta} \qquad (t < 0) \tag{91}$$

$$\chi_T \propto |t|^{-\gamma} \qquad (\psi_> = 0; \psi_< = \bar{\psi}) \tag{92}$$

$$h \propto |\psi|^{\delta} \qquad (t = 0) \tag{93}$$

$$\xi \propto |t|^{-\nu} \tag{94}$$

Broadly speaking, α characterizes the thermal anomaly and γ the mechanical, magnetic, etc. instability, at the transition. For the liquid-gas critical point $\chi_T = \kappa_T$, and γ quantifies the divergence of the compressibility. The shape of the critical isotherm $h(\psi)$ is characterized by δ. The notation $|t|$ indicates that both positive and negative t are to be considered, and that the singular behavior above and below the transition is quite symmetric (except for numerical factors). The emerging order parameter that characterizes the new quality emerging "below" the critical point is quantified by β (defined only for $t < 0$), and its spatial correlations by the correlation length ξ with exponent ν. A *spontaneous symmetry breaking* is often associated with the transition; e.g., a spontaneously emerging (self-)magnetization clearly has to point in some definite but *a priori* arbitrary direction, and domain walls break homogeneity but can come up anywhere.

There are two perplexing experimental facts about critical phenomena:

1. Their *universality*, meaning that phenomenologically very different systems such as magnets and fluids may share the identical critical behavior (the same critical exponents), which allows to group the wide variety of materials and other physical systems into a small number of universality classes.

2. Up to rare exceptions, the *critical exponents are irrational numbers* related by so-called *exponent relations*.

Landau–Ginzburg theory

A general way to deal with small order parameter fluctuations ψ at a critical point is via Landau–Ginzburg (LG) theory, which is a DFT of the weakly-inhomogeneous type discussed above. It assumes the existence of a Taylor expansion of the intrinsic free energy (the Legendre transform of the grand potential in the order parameter and its conjugate field) in the small deviations $\psi(\mathbf{r})$ ("order parameter"), $t \equiv T/T_c - 1$ ("critical parameter") and $h(\mathbf{r})$ ("conjugate field"), around the critical point:

$$L_G[\psi(\mathbf{r}), h(\mathbf{r})] = n_c k_B T_c \int_V \mathrm{d}\mathbf{r}\, \mathcal{L}_G \,, \quad \mathcal{L}_G \equiv \frac{\ell^2}{2}(\nabla\psi)^2 + \frac{t}{2}\,\psi^2 + \frac{g}{4}\,\psi^4 - h\psi \,. \tag{95}$$

For a fluid, ψ and h correspond to the distance from the critical density and the distance from the critical chemical potential, for a magnet to magnetization and external magnetic field, respectively. The fields are understood to be normalized to the critical values n_c and T_c of the density and temperature, in order to make them dimensionless. In fact, the whole free energy functional is made dimensionless by the factor $n_c k_B T_c$ and by ℓ being a length and g some number. The generic form of the Landau functional can be written down for any system of interest without a precise knowledge of the direct correlation function, simply on the basis of the dimensionalities of the order parameter and the embedding space, and considering only terms in the Taylor expansion that are allowed by symmetry[21]. For example, for an up-down symmetric magnet no odd powers of the magnetization are allowed.

To second order in the order parameter field, the expansion has two contributions. The factor in front of the quadratic term $\propto \psi^2$ is the leading term of the inverse pair correlation $G^{-1}(\mathbf{r}, \mathbf{r}')$ for $q \to 0$, i.e., for a homogeneous system it is a response coefficient (the inverse isothermal compressibility κ_T^{-1} in case of a fluid), and in accord with Eq. (89) it is denoted by t (the dimensionless critical temperature parameter). In the spirit of hydrodynamic theories, long-wavelength spatial structure (to order q^2) is encoded in the gradient term, which suggests that the length ℓ should be on the order of the range of the direct correlation function (roughly the interaction range[22]).

So, Landau–Ginzburg theory is clearly a DFT of the sort devised above to deal with weak long-wavelength fluctuations about a reference state, namely the critical state. Upon minimizing L_G (i.e. going over from the Legendre to a Legendre–Fenchel transform), it correctly predicts the emergence of order below T_c. And it suggests that thermodynamic behavior may be classified into universality classes according to the topology of the embedding space and the order parameter field and according to fundamental symmetries. So it accounts for observation (1) from the preceding paragraph.

However, as already pointed out, such a "hydrodynamic" theory cannot predict non-mean-field values for the critical exponents. So the theory cannot account for observation (2) from the preceding paragraph. The failure of the theory is not merely a quantitative problem, which could be overcome by a slightly improved model. It is a failure in principle, due to the basic assumption that the free energy (or the direct correlation function) may be expanded in a Taylor series in the small parameters. It also would not help to simply replace the factor of t in the Landau functional by t^γ, as this would be inconsistent with the thermodynamically required exponent relations discussed next.

---

[21] This may sound familiar from the standard model of particle physics — a Landau theory.

[22] Do not worry about the precise values of any of the factors for any particular system, it would be against the spirit of the whole approach, which aims at universal aspects.

Scaling hypothesis and exponent relations

As pointed out above, a major shortcoming of Landau theory and density-functional theories, in general, is their failure to properly predict the non-trivial behavior at continuous phase transitions and critical points. It is also not easy to repair the theory by hand, since the critical exponents are known (phenomenologically and thermodynamically) to obey a set of nontrivial exponent relations, i.e., they are not independent. It turns out that this observation can be rationalized by a postulate known as the scaling hypothesis. It consists in the requirement of a generalized homogeneity relation for the Landau free energy $\mathcal{L}$, or rather its *singular part* $\mathcal{L}_s$,

$$\lambda \mathcal{L}_s(t,h) = \mathcal{L}_s(\lambda^{a_t} t, \lambda^{a_h} h) \; , \tag{96}$$

where λ is a positive scale factor and a_t and a_h are real numbers. In particular, choosing $\lambda = |t|^{-1/a_t}$, Eq. (96) reduces to

$$\mathcal{L}_s(t,h) = |t|^{1/a_t} \mathcal{L}_s(\pm 1, h/|t|^{a_h/a_t}) \equiv |t|^{1/a_t} \mathcal{L}_\pm(h/|t|^{a_h/a_t}) \; . \tag{97}$$

Landau's analyticity assumption for $\mathcal{L}$ is now replaced by regularity requirements for the *scaling function* $\mathcal{L}_\pm$ and its first and second derivatives at the origin, complemented by the asymptotic condition $\mathcal{L}'_\pm \equiv \partial_x \mathcal{L}_\pm \overset{x\to\infty}{\sim} -x^{1/\delta}$ (to reproduce the critical isotherm). Then Eq. (97) immediately produces the desired power-law behavior for $t \to 0$:

$$\begin{aligned}
C_V &\propto -T\partial_t^2 \mathcal{L}_s(t,h=0) \propto |t|^{1/a_t - 2} \mathcal{L}_\pm(0) && (98)\\
\psi &\propto -\left.\partial_h \mathcal{L}_s\right|_{h=0} \propto |t|^{(1-a_h)/a_t} \mathcal{L}'_\pm(0) && (99)\\
\chi_T &\propto -\left.\partial_h^2 \mathcal{L}_s\right|_{h=0} \propto |t|^{(1-2a_h)/a_t} \mathcal{L}''_\pm(0) && (100)\\
\psi(h) &\propto -\partial_h \mathcal{L}_s \sim -|t|^{(1-a_h)/a_t} \mathcal{L}'_\pm(h/|t|^{a_h/a_t}) && (101)\\
&\propto |t|^{1/a_t - a_h/a_t(1+1/\delta)} h^{1/\delta} \qquad (t \to 0) \; .
\end{aligned}$$

Comparison with the definitions in Eqs. (90)-(94) moreover yields

$$\begin{aligned}
\alpha &= 2 - 1/a_t && (102)\\
\beta &= (1-a_h)/a_t && (103)\\
\gamma &= (2a_h - 1)/a_t && (104)\\
\delta &= a_h/(1-a_h) \; . && (105)
\end{aligned}$$

The fact that two independent exponents create a substantial part of the Greek alphabet immediately explains the existence of exponent relations, such as $2-\alpha = 2\beta+\gamma$, which were historically postulated on thermodynamic grounds (exercises), and originally motivated the scaling hypothesis, in the first place. Also note that equipartition implies[23] that for $T \approx T_c$ correlation volumes ξ^d have a free energy

---

[23] Equipartition means that each degree of freedom carries $k_B T/2$ of thermal energy in its fluctuations, as corroborated by an explicit calculation for the LG energy, below.

$L_s(V=\xi^d)$ equal to the thermal energy k_BT_c. In other words,

$$n_c k_B T_c \int_{\xi^d} \mathrm{d}\mathbf{r}\, \mathcal{L}_s \simeq n_c k_B T_c \xi^d \mathcal{L}_s \simeq k_B T_c \;, \tag{106}$$

so that the free-energy density scales like $\mathcal{L}_s(t,h=0) \propto n_c^{-1}\xi^{-d}$. Since the critical density n_c is just some constant (on the order of ℓ^{-d} with ℓ being an estimate for the range of the pair interaction), the correlation length diverges at the critical point such that $\xi^{-d} \sim n_c|t|^{1/a_t}$, according to Eq. (97). With, Eqs. (94), (102), the so-called hyperscaling relation $d\nu = 2-\alpha$ follows, which depends on the space dimension d. It says that the values of the exponents are not only mutually dependent but that there is a distinct set of such exponents for each d.

The Ginzburg criterion and the upper and lower critical dimension

The above scaling hypothesis is an entirely phenomenological approach to critical phenomena. It postulates a specific scaling form of the free energy but does not give a clue how it might arise from a microscopic description. One can press Landau–Ginzburg theory a bit harder to better understand the reason for its breakdown, and from there develop an idea about the origin of the scaling in Eq. (96). Namely, as a DFT for small fluctuations around the critical reference state, it is expected to fail when the (relative) fluctuations become large, which, because of the vanishing/diverging transport coefficients, actually happens right at the critical point. As a reference for what "large" means, one can take the equilibrium value $\bar\psi$ of the order parameter for $t<0$ and $h=0$, which follows from minimizing the (homogeneous) Landau free energy density, namely

$$\mathcal{L} = \frac{t}{2}\psi^2 + \frac{g}{4}\psi^4 \quad \overset{\partial_\psi \mathcal{L}=0}{\Rightarrow} \quad \bar\psi = \pm\sqrt{-t/g} \;. \tag{107}$$

Note that this implies the prediction $\beta = 1/2$. The theory's estimate for the size of the critical parameter region around the critical point, where this prediction becomes inaccurate, is known as the Ginzburg criterion. The idea is the following. When the quadratic term vanishes at the critical point, the task of confining the order parameter fluctuations is largely left to the gradient term, since the ψ^4–valley is very flat. Using dimensional analysis, the relevant integration range and the ∇ are estimatd by the characteristic correlation length ξ, i.e. $\int\mathrm{d}\mathbf{r} \to \xi^d$, $\nabla \to \xi^{-1}$. One now repeats the above argument with equipartition for the fluctuations in correlation volume elements ξ^d (as effective collective degrees of freedom near the critical point), but this time for the explicit model provided by the LG free energy. This yields

$$n_c k_B T_c \int_{\xi^3} \mathrm{d}\mathbf{r}\, \ell^2 (\nabla\psi)^2 \simeq n_c k_B T_c \xi^{d-2}\ell^2 \delta\psi^2 \simeq k_B T_c \;. \tag{108}$$

Here $\delta\psi \equiv \psi - \bar{\psi}$ is the order-parameter fluctuation around the equilibrium state. Its strength follows as

$$\delta\psi^2 \simeq \ell^{-d} n_c^{-1} (\xi/\ell)^{2-d} \simeq |t|^{(d-2)\nu} \overset{\text{LG}}{=} |t|^{(d-2)/2} \; . \tag{109}$$

The last expression employs the LG prediction $\nu = 1/2$ corroborated in Eq. (114), below. Now, compare this with the (squared) absolute value $\bar{\psi}^2 = |t|/g$ of the order parameter. Both $\delta\psi$ and $\bar{\psi}$ are functions of t, and one would expect the Ginzburg theory to be trustworthy in a range of (small) t where the fluctuations $\delta\psi$ do not exceed $\bar{\psi}$ in magnitude:

$$\delta\psi^2/\bar{\psi}^2 \simeq \ell^{-d} n_c^{-1} g |t|^{(d-4)/2} \lesssim 1 \quad \Rightarrow \quad |t|^{(d-4)/2} \lesssim n_c \ell^d / g \; . \tag{110}$$

This is the so-called *Ginzburg criterion.* It shows that $d = d_u = 4$ plays the role of an upper critical dimension. For $d > d_u$ fluctuations vanish for $t \to 0$ and Landau theory should hold close to (and at) the critical point. In contrast, for $d < d_u$, Eq. (110) defines a critical region $|t| \lesssim (g/\ell^d n_c)^{2/(4-d)}$ around the critical point where Landau theory fails. In brief: “murder at the critical point — macroscopic order parameter fluctuations kill Landau theory.” Mean-field theory thus breaks down close to the critical point, while further away from criticality (albeit not too far for a Taylor-expanded Landau free energy) one can trust its predictions (sketch). As an exception to the rule, “classical” superconductors do obey mean-field theory, since their critical region is very small due to the large off-critical correlation length $\ell \gg n_c^{-1/3}$ due to the strongly delocalized Cooper pairs.

The above estimate of the order parameter fluctuations $\delta\psi$ in Eq. (109) moreover suggests that these become large in absolute terms (not only relative to $\bar{\psi}$), below $d \lesssim d_l = 2$. This hints at an even more severe breakdown of the theory below a *lower critical dimension* d_l. This suspicion, and actually the whole above discussion, is corroborated by a straightforward diagonalization of L_G in the harmonic approximation (exercises), by means of the Fourier modes

$$\psi_{\mathbf{q}} \equiv \frac{1}{\sqrt{V}} \int_V \mathrm{d}\mathbf{r}\, \psi(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} \qquad \psi(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{q}} \psi_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} \; , \tag{111}$$

$$L_G = n_c k_B T_c \frac{1}{2} \sum_{\mathbf{q}} (t + q^2 \ell^2) |\psi_{\mathbf{q}}|^2 \; . \tag{112}$$

For simplicity, $T \gtrsim T_c$ ($t > 0$) and $h = 0$ shall be assumed. Invoking equipartition for the eigenmodes $\psi_{\mathbf{q}}$, the average strength of the order parameter fluctuations follows from

$$n_c k_B T_c (t + q^2 \ell^2) \langle |\psi_{\mathbf{q}}|^2 \rangle = k_B T_c \; . \tag{113}$$

Observe that, while the average order parameter $\langle \psi \rangle$ vanishes for $T > T_c$, its squared fluctuations have non-vanishing mode amplitudes

$$n_c \ell^2 \langle |\psi_{\mathbf{q}}|^2 \rangle = \frac{1}{\xi^{-2} + q^2} \; . \tag{114}$$

Here the parameter combination $\xi \equiv \ell/\sqrt{|t|}$ is identified with the correlation length of order parameter fluctuations, which corroborates the LG value $\nu = 1/2$ employed in Eq. (109). The correlation length controls the spatial decay of correlations, which becomes more evident from the Fourier back transform, the order parameter *correlation function* in real space,

$$h(\mathbf{r}) = \langle \psi(\mathbf{r})\psi(0)\rangle = \mathrm{FT}^{-1}\frac{\langle|\psi_{\mathbf{q}}|^2\rangle}{\sqrt{V}} \propto \begin{cases} \dfrac{e^{-r/\xi}}{r^{(d-1)/2}} & (r \to \infty) \\ r^{(2-d)} & (\xi \to \infty) \,. \end{cases} \tag{115}$$

The average of the product of two values of the order parameter at two points a distance r apart does not vanish unless $r \gg \xi$, although $\langle\psi(\mathbf{r})\rangle$ is zero for $T > T_c$.

From the last expression in Eq. (115), one can again read off the lower critical space dimension $d_\mathrm{l} = 2$. Correlations should decay with distance, not grow indefinitely. The more general (perplexing) observation that hydrodynamic fluctuations can cause the complete breakdown of hydrodynamic theories below a lower critical dimension d_l (not only a lack of precision, as below the upper critical dimension) has been called hydrodynamic suicide[24]. It is not an artifact of the theory but due to a real physical effect. Intuitively, due to the reduced number of constraining neighbors, long-wavelength hydrodynamic excitations trying to restore the broken symmetries (also called Goldstone modes) become all pervasive at and below the lower critical dimension $d_\mathrm{l} = 2$ at any $T > 0$. That this prevents the phase ordering phenomena observed in higher dimensions (although not phase separation as such) for continuous order parameters in $d \leq 2$, if the interactions are not themselves long-ranged, has been proved in great generality ("Mermin–Wagner theorem"). It implies that there are, in the thermodynamic limit, no (proper) one- or two-dimensional crystals. This hints at the fact that the physics of membranes and polymers is dominated by fluctuations[25]. Phase transitions exactly at $d = d_\mathrm{l}$ (much discussed by Kosterlitz, Thouless, Nelson and Halperin) are quite subtle and have kept statistical physicists busy for decades.

There is another important subtlety related to the second Eq. (115). It is a prediction for the cirtical behavior inside the critical region, where one should actually not trust the prediction of the Ginzburg theory. Precise measurements have indeed revealed small systematic deviations from the Ginzburg form for $r \ll \xi$ in $d < 4$ space dimensions. They have (first phenomenologically) been rationalized by introducing a so-called "anomalous dimension" η, writing

$$n_c \ell^2 \langle|\psi_{\mathbf{q}}|^2\rangle \propto q^{-2+\eta}\,, \qquad \langle\psi(\mathbf{r})\psi(0)\rangle \propto 1/r^{d-2+\eta}\,, \qquad \xi \to \infty \,. \tag{116}$$

The name derives from the perplexing observation that this relation apparently violates ordinary dimensional analysis. It implies that there must be a subtle

---

[24] W. Brenig, *Statistical Theory of Heat, Nonequilibrium Phenomena.*

[25] In these cases, fluctuations are even stronger than suggested by the theorem, which does not consider embedding space dimensions exceeding the dimension of the system manifold, but states that, even without them, crystaline order would be destroyed by thermal noise.

hidden dependence on the microscopic length scale ℓ with an odd irrational exponent η, even — or rather in particular — at the critical point, where $\ell/\xi \to 0$. This is a consequence of the strong coupling of fluctuations on all scales (large and small) at the critical point. As a side remark, Eq. (116) implies the exponent relation $\gamma = \nu(2 - \eta)$. It follows from the compressibility sum rule that relates the order-parameter fluctuations in the form of the integral over $h(\mathbf{r})$ to the corresponding response coefficient (the isothermal compressibility). This relation clearly is not limited to gases or fluids. For a magnet, one simply has to replace particle number fluctuations $\langle N^2 \rangle$ by spin number (or magnetization) fluctuations and the isothermal compressibility κ_T by the magnetic susceptibility χ_T to arrive at the same conclusion.

4.2 Learning from toy models: the Ising chain

The discussion in the previous section has revealed that the DFT-type or "hydrodynamic" approach of the Landau–Ginzburg theory breaks down near the critical point and in the physically interesting space dimensions below an upper critical dimension ($d_{\mathrm{u}} = 4$). It thereby pointed at the important role of fluctuations for the critical behavior. Below the critical dimension, a critical region develops, in which the fluctuations get so large as to invalidate Landau–Ginzburg theory. A good strategy to gain further insight into the origin of the non-trivial critical behavior inside the critical region, is to analyze toy models and try to extract a general strategy for the study of critical behavior. An interesting simple model is the so-called Ising chain. Unfortunately, it comes with two disadvantages that make it non-generic. First, the Ising model has a discrete (not a continuous) state space, so that its thermal excitations are not hydrodynamic Goldstone modes (namely spin waves, as for most real magnets) but defects. This is sad, but not too bad, because defects are interesting in their own right. Also, as a consequence, the lower critical dimensions is $d_{\mathrm{l}} = 1$ instead of 2, which is nice, since it allows to have a phase transition even for the simple one-dimensional chain. The drawback is that a phase transition at the lower critical dimension is a non-generic example. (It would be nicer to study the Ising model in two spatial dimensions, which is however considerably more complicated to solve.)

Historically, Onsager's 1944 solution of the two-dimensional Ising model with nearest neighbor interactions settled debates about the possibility of non mean-field values for critical exponents, and whether these can be deduced from a partition sum, at all. In the 1950's, the theory of Yang-Lee and Fisher zeros finally elucidated the general mathematical mechanism how a singular free energy may emerge from a sum of infinitely many positive analytical Boltzmann factors. A more general understanding of critical behavior and its universality as a consequence of self-similar fluctuations emerged with the development of the renormalization group theory in the 1970s by Fisher, Kadanoff, Wilson (who received the Nobel prize in 1982) and others.

The Ising model

The Ising model consists of a regular lattice of N binary "spin" variables s_i, which can take the values $s_i = \pm 1$ (or $s_i = \pm 1/2$, $s_i = 1, 0$). Depending on the application one has in mind, these variables can e.g. represent constrained magnetic moments that can only point up or down, or the presence or absence of atoms in a lattice model of a gas, fluid or solid, votes in an election, firing states of neurons, etc. etc. Depending on the type and dimensionality of the lattice and the spin interactions, the Ising model may thus serve as an idealized minimal model for a large variety of complex physical, biological, or even social systems encountered in the real world.

In presence of an external field h, the Hamiltonian reads

$$H(\{s_i\}) = -J \sum_{\text{ip}} s_i s_j - h \sum_i s_i \tag{117}$$

Ferromagnetic and anti-ferromagnetic interactions can be realized by choosing a positive or negative J, respectively. (More generally, one may consider a spin glass with a non-thermal quenched distribution of interactions J_{ij}, if one is after a truly complex non-equilibrium model.) The interaction sum runs over all N_{ip} interacting pairs ("ip"). Two very useful idealizations are provided by the extreme choices that the pair interactions are either (1) van-der-Waals type *infinite-range interactions* ($\sum_{\text{ip}} \to \sum_{i<j}$, $N_{\text{ip}} = N(N-1)/2$, with $J \to I/N$ to make H extensive), which produces mean-field behavior, and (2) *nearest-neighbor interactions* ($\sum_{\text{ip}} \to \sum_{\langle ij \rangle}$, $N_{\text{ip}} = Nq/2$, with q being the coordination number, i.e. the number of nearest neighbors per lattice site), which produces non-trivial critical behavior below 4 dimensions. Writing $s_i = \sigma + (s_i - \sigma)$ in terms of the mean spin magnetization $\sigma = \langle s_i \rangle$ plus fluctuations $s_i - \sigma$ around it, and neglecting correlations of fluctuations according to

$$s_i s_j = -\sigma^2 + \sigma(s_i + s_j) + (s_i - \sigma)(s_j - \sigma) \approx -\sigma^2 + \sigma(s_i + s_j) \ , \tag{118}$$

amounts to treating the spins as independently exposed to the self-generated mean field $Jq\sigma$, i.e. to the mean-field approximation, which becomes exact if the number of neighbors $q \simeq N - 1$ (or equivalently the interaction range) becomes extensive, which requires $J \to I/N$ with some fixed I to maintain an extensive energy. If one calculates the magnetization $\sigma = \langle s_i \rangle$ with the mean-field Hamiltonian $H = JN\sigma^2 - J\sigma \sum s_i + s_j$, one recovers the usual self-consistency equation for the weakly-inhomogeneous density-functional theory, in random-phase approximation (which is thereby also seen to be exact, in this case).

Defects and the low-temperature expansion

In the most interesting case of low dimensions and finite-range interactions, fluctuations change the mean-field picture, substantially. A general strategy to address these effects is the low temperature expansion, which starts from the homogeneous ground state for $T = 0$ and counts the number of defects induced by a slight temperature increase. Consider, for simplicity, the 1-dimensional Ising spin chain with nearest-neighbor coupling, consisting of N spins with $N - 1 \approx N$ bonds. In the ground state all spins are aligned. Adding some thermal energy to topple some spins, one creates so-called *defects*, namely "broken bonds" between pairs of misaligned spins. The creation of a defect costs an energy $2J$. An individual bond between spins can either be satisfied or broken, corresponding to the Boltzmann factors $e^{\beta J}$ and $e^{-\beta J}$, and to the partition sum $Z = e^{\beta J} + e^{-\beta J}$ for each bond. The fraction of broken bonds in equilibrium is then simply[26] $(1 + e^{2\beta J})^{-1}$. In other words, to completely freeze out the defects, one has to cool to absolute zero.

Arguing along these lines, one can even write down the full partition sum, by ordering the its terms according to the the number N_b of broken bonds or defects (or according to the associated excitation energy $2JN_b$):

$$Z^{(1\mathrm{d})} = 2e^{N\beta J} \sum_{N_b=0}^{N} \tbinom{N}{N_b} e^{-2N_b\beta J} = 2e^{N\beta J}[1 + Ne^{-2\beta J} + (N/2)(N-1)e^{-4\beta J} + \dots]$$

The leading term is the Boltzmann factor of the twofold degenerate ground state (all spins up or down), the other terms correspond to one defect, two defects, etc., which can occur anywhere along the chain, i.e. in N_b out of N places. This form of the partition sum is particularly useful at low temperatures, where accurate approximations are obtained by truncating the sum after a finite number of terms. Note that the strategy behind this solution is that the attention is shifted from spins with two configurations (up/down) that are mutually interacting to "bonds" with two configurations (satisfied/broken) that are mutually independent.

Another way of looking at defects is the following. Single defects can wander about along the chain at no cost like gas particles. Low temperatures, $\beta J \gg 1$, imply a small number N_b of broken bonds, corresponding to a defect fraction $\phi \equiv N_\mathrm{b}/N \ll 1$. Hence the probability ϕ^2 that two defects meet (and annihilate) is very small, so that they can be treated as an ideal gas with entropy $S = k_B N_\mathrm{b} \ln(T^{1/2}/\phi)$ and defect free energy[27]

$$F = U - TS = N_\mathrm{b}\left[2J - k_B T \ln(T^{1/2}/\phi)\right] \ . \tag{119}$$

Whenever the term in brackets is negative, which is certainly the case for $\phi \to 0$ at $T > 0$, the free energy can be lowered by increasing the number N_b of defects. That is, if you had prepared the system in a uniform phase (e.g. all spins

---

[26] With periodic boundary conditions, elementary defects correspond to pairs of broken bonds, hence their fraction is $2(1 + e^{2\beta J})^{-2}$.

[27] Unlike the kinetic energy, which is often dismissed as an irrelevant gauge term, the energetic cost for defect creation matters (as does the rest mass in a relativistic gas).

down), the chain would spontaneously split up into up- and down-domains by a spontaneous creation of defects. For $\phi \to 0$, the competition between energy and entropy in the free energy is dominated by entropy. The opposite is expected in phase separation or phase ordering processes, in which small domains grow and coarsen until they occupy a macroscopic fraction of the sample homogeneously, as favored by energy. This process reduces the number of defects. Note that the limits $T \to 0$ ($\beta J \to \infty$) and $N \to \infty$ do not interchange: any finite chain is ordered at $T = 0$, while an infinite chain is disordered at any finite T). This is indicative of a non-analytic structure at $T = 0$, $N = \infty$, which is studied in some detail further below.

In one dimension, defects thus proliferate at any finite temperature $T > 0$. In higher dimensions, defects come as lines and (hyper-)surfaces separating domains of spins with opposite orientation. Because of their energetic cost, these have the tendency to contract and freeze out the enclosed minority orientation, at low temperatures. As a consequence, spins of one orientation will start to dominate over those with the other, thereby giving rise to a spontaneous magnetization, below a finite critical temperature $T_c > 0$, consistent with a lower critical dimension $d_\mathrm{l} = 1$, as first argued by R. Peierls in 1935.

Transfer matrix for the Ising chain

A nice way to solve the 1-dimensional Ising chain with nearest neighbor interactions and periodic boundary conditions is the transfer matrix method. Its starting point is the observation that the partition sum

$$Z = \sum_{\{s_i=\pm 1\}} e^{-\beta H} = \sum_{\{s_i=\pm 1\}} e^{\beta \sum_i J s_i s_{i+1} + h s_i} \ . \tag{120}$$

decays into a product of structurally identical terms

$$Z = \sum_{\{s_i=\pm 1\}} \prod_i e^{\beta J s_i s_{i+1} + \beta h (s_i + s_{i+1})/2} \ . \tag{121}$$

This can be rewritten using an operator $\hat{T}$ with matrix elements

$$\langle s_i | \hat{T} | s_{i+1} \rangle = \exp[\beta J s_i s_{i+1} + \beta h (s_i + s_{i+1})/2] \tag{122}$$

and matrix representation

$$\begin{pmatrix} \exp[\beta(J+h)] & \exp[-\beta J] \\ \exp[-\beta J] & \exp[\beta(J-h)] \end{pmatrix} \qquad (\text{“transfer matrix”}) \tag{123}$$

in the spin space with basis $\{|+\rangle, |-\rangle\}$ for $s_i = \pm 1$:

$$Z = \sum_{\{s_i=\pm 1\}} \langle s_1 | \hat{T} | s_2 \rangle \langle s_2 | \hat{T} | s_3 \rangle \dots \langle s_N | \hat{T} | s_1 \rangle \ . \tag{124}$$

Using the completeness of the spin basis, one has

$$Z = \sum_{s_1=\pm 1} \langle s_1 | \hat{T}^N | s_1 \rangle = \mathrm{tr} \hat{T}^N = \lambda_+^N + \lambda_-^N \ , \tag{125}$$

where λ_+ and λ_- denote the transfer matrix eigenvalues

$$e^{\beta J}\big[\cosh \beta h \pm (e^{-4\beta J} + \sinh^2 \beta h)^{1/2}\big] \sim \begin{cases} e^{\beta J} \pm e^{-\beta J} & (\beta h = 0) \\ 2\cosh \beta h \, , \, 0 & (\beta J = 0) \\ e^{\beta J \pm \beta |h|} & (\beta J \to \infty) \end{cases} \ . \tag{126}$$

Emergence of singularities in the double limit $T \to 0$, $N \to \infty$

Observe that for finite βJ and βh the transfer matrix is symmetric and has positive finite entries, so there is a non-degenerate largest eigenvalue $\lambda_+ > \lambda_-$, which is an analytical function of βJ and βh. This is the Frobenius–Perron theorem, which is however potentially invalidated if matrix entries diverge or vanish in the limits taken in Eq. (126). As long as the eigenvalues are non-degenerate, the larger eigenvalue λ_+ will dominate the partition sum in the thermodynamic limit ($N \to \infty$), corresponding to the free energy per spin

$$\beta f = \lim_{N\to\infty} \frac{\beta F}{N} = -\ln \lambda_+ = \begin{cases} -\ln[2\cosh \beta J] - (\beta h e^{\beta J})^2/2 & (\beta h \to 0) \\ -\ln[2\cosh \beta h] & (\beta J = 0) \\ -\beta J - |\beta h| & (\beta J \to \infty) \end{cases} \ .$$

The low-temperature limit $\beta J \to \infty$, in which matrix entries diverge or vanish, that does indeed produce the anticipated non-analytical behavior characteristic of a phase transition. Below, it is verified that the result stays valid if the order of the limits is reversed. In the high-temperature limit without external field ($\beta J \to 0$, $h = 0$), F reduces to $-TS$ with $S = k_B N \ln 2$ the entropy of the up-down degeneracy for each spin in a paramagnet. Further, taking derivatives of f with respect to h, gives the site magnetization

$$\sigma \equiv \langle s_i \rangle = -\partial_{\beta h} \beta f = \frac{\sinh \beta h}{\sqrt{\sinh^2 \beta h + e^{-4\beta J}}} \sim \begin{cases} \beta h e^{2\beta J} & (\beta h \to 0) \\ \tanh \beta h & (\beta J \to 0) \\ \mathrm{sgn}(\beta h) & (\beta J \to \infty) \end{cases} \tag{127}$$

and the susceptibility

$$\chi = \partial_h \langle \sigma \rangle|_{h=0} = \beta e^{2\beta J} \ . \tag{128}$$

The latter reduces to the well-known Curie law $\chi \simeq \beta$ (the constitutive equation for a paramagnet) at high temperature ($\beta J \to 0$), but is enhanced compared to the latter by a strange (non-powerlaw) divergent coefficient[28] at low temperature

---

[28] This semiconductor-type “Schottky anomaly” is due to the finite energy gap between the ground state ($T = 0$) and its excitations (no Goldstone modes with arbitrarily small energy).

($\beta J \to \infty$). It is suggestive of the unusual critical parameter $t = e^{-\beta J}$ and the critical exponent $\gamma = 2$, a point that is elaborated further, below.

Now, reversing the order of limits, i.e. keeping N finite and taking the limit $\beta J \to \infty$ first, and using $(1 + x/N)^N \sim e^x$, Eqs. (125) & (126) give for small βh

$$Z = e^{N\beta J} \left[(1 + |\beta h|)^N + (1 - |\beta h|)^N\right] = e^{N\beta J} 2\cosh(N\beta h) \; , \tag{129}$$

hence

$$\beta f \equiv \frac{\beta F}{N} = -\beta J - \frac{1}{N} \ln[2\cosh(N\beta h)] \overset{N\to\infty}{\simeq} -\beta J - |\beta h| \; . \tag{130}$$

While the ultimate result in the double limit $\beta J,\ N \to \infty$ is the same as found above by taking the limit in the reverse order, it is interesting to see also the sub-leading behavior and to analyze the non-extensive contributions for finite N. The argument of the cosh is no longer the energy h of a single spin in an external field, as in a paramagnet, but instead the energy Nh of N aligned spins in the field h. This is reflected in the site magnetization

$$\sigma = -\partial_h f = \tanh N\beta h \sim \operatorname{sgn}(\beta h) \, , \tag{131}$$

and in the susceptibility

$$\chi = -\partial_h^2 f = \frac{N\beta}{\cosh^2 N\beta h} \sim 2\beta\delta(\beta h) = 2\delta(h) \; . \tag{132}$$

They both clearly display how the singularity develops (for $\beta J \to \infty$) from smooth functions as a consequence of spontaneous collective behavior, upon taking the thermodynamic limit $N \to \infty$. For $T \to 0$ all spins align and are turned by an infinitesimally small field like a single giant spin $\mathcal{S} = Ns_i$, hence $\chi = \beta N \to \infty$, for $h = 0$. In contrast, once the giant spin has been oriented by an infinitesimal field, increasing h has no further effect, so that $\chi = 0$ for $h \neq 0$.

The above discussion of the order of the two limits $\beta J \to \infty$ and $N \to \infty$ corroborates the physical intuition that, for any finite $J \neq 0$, a finite chain will order at $T = 0$, while an infinite chain will always be disordered at any $T > 0$. It also provides a straightforward interpretation of Eq. (129), which is simply the partition sum of a paramagnet consisting of a single spin of length N (shifted by the ground state energy $-JN$), as could have been guessed without calculation.

The consequence of the formation of a giant spin can also be seen in the behavior of the specific heat $c_h = -T\partial_T^2 f$. Starting from Eq. (130), which corresponds to the taking the limit $N \to \infty$ after $\beta J \to \infty$, reveals a tiny specific heat

$$\frac{c_h}{k_B} = \frac{N(\beta h)^2}{\cosh^2 N\beta h} \sim \frac{\pi^2}{12N}[\delta^+_{\beta h} + \delta^-_{\beta h}] \; . \tag{133}$$

The symbolic notation $\delta^\pm$ refers to a Kronecker$-\delta$ centered around slightly positive or negative arguments. As all spins align into a giant spin $\mathcal{S} = Ns_i$ in the low-temperature limit, the free energy becomes simply the *ground state energy* $-J - |h|$. The only remaining entropy is the non-extensive contribution $k_B \ln 2$, as seen by setting $h = 0$ in Eq. (130) before taking the thermodynamic limit $N \to \infty$. It corresponds to the two-fold ground state degeneracy, i.e. to turning the giant spin. For this to happen the field has to be switched off, since otherwise $\mathcal{S}$ is firmly aligned along h. Accordingly, only a small, non-extensive amount $\approx k_B$ of heat can be taken up, altogether, where the specific heat vanishes in the thermodynamic limit. Yet, the limit in Eq. (133) is subtle and interesting. To reveal the specific heat, the ergodicity of the ensemble has to be broken artificially by an infinitesimally small field ($|h| > 0$) to single out the up or the down component. (This is what typically would automatically happen in a real experiment.) The giant spin can then explore the other half of the ensemble after the field is weakened enough ($h \to 0$) and some heat is provided. This doubles the accessible phase space volume, which manifests itself as a (tiny) specific heat. In contrast, setting $\beta h = 0$ from the outset would have allowed both sub-components of the phase-space to remain equally populated. This would have killed the specific heat entirely, as it formally corresponds to $T = \infty$, meaning that the system cannot be heated up.

Recovering the scaling hypothesis and finite-size scaling

To recover the scaling hypothesis from the above exact solution for the free energy of the Ising chain, consider first again the logarithm of the largest eigenvalue Eq. (126), keeping only the leading order terms in the critical parameters $e^{-\beta J}$ and βh for large βJ, to find

$$\beta f \sim -\beta J - e^{-2\beta J}\sqrt{1 + (\beta h e^{2\beta J})^2}\ . \tag{134}$$

The singular part f_s of the free energy f per spin, which is up to a division with the lattice constant the same as the free energy density $\mathfrak{f}$, is indeed of the expected scaling form

$$\beta f_s = \beta f + \beta J \sim -\hat{\xi}^{-1}\varphi(\beta h \hat{\xi}) = -t^{\nu}\varphi(\beta h / t^{\nu}) = -t^{2-\alpha}\varphi(\beta h / t^{\gamma+\beta})\ . \tag{135}$$

Alternatively, starting from the expression for large (but finite) N, βJ and $\beta h = 0$, the free energy follows from Eq. (152) as

$$\begin{aligned}\beta f_s \sim& -\frac{1}{N}\ln[(2\cosh(N e^{-2\beta J})]) \\ \sim& -e^{-2\beta J}\big[N e^{-2\beta J} - N^3 e^{-6\beta J}/6 + \dots\big]/2 \\ \sim& -\hat{\xi}^{-1}\phi(N/\hat{\xi}) \qquad\qquad (\beta J \to \infty\,,\ N \to \infty)\end{aligned} \tag{136}$$

where $\phi = \mathcal{O}(1)$ should hold for large and small arguments, to reproduce the critical behavior, and because the free energy density must be intensive. From the specific form of the free energy in Eq. (136), which is called a *finite-size scaling* relation, one reads off that any finite chain already becomes fully critical at a length-dependent finite distance from the critical point, namely as soon as $\hat{\xi}$ exceeds N, so that $\phi(N/\hat{\xi}) \sim \phi(0)$ saturates. Or, turned the other way round, one can find the exponent ν from comparing the free energy of finite chains of different lengths N. By tuning the temperature such that the singular part of the free energy density vanishes like $f_s \propto 1/N$, one knows that one has achieved $N/\hat{\xi} =$constant, so that one can infer the divergence of the correlation length with temperature. The method is not limited to the free energy, of course, and is in fact a standard trick used in computer simulations to extract precise values of critical exponents for systems of finite size.

4.3 Renormalization and Yang–Lee zeros

Remember that, according to the Ginzburg criterion, the critical region shrinks to zero above the upper critical dimension, $d_{\mathrm{u}} = 4$, where DFT theories remain valid at the critical point. In that sense, a dimension $d > d_{\mathrm{u}}$ is as good as infinitely many dimensions. A successful general strategy to gain access to the true critical exponents in dimensions $d < d_{\mathrm{u}}$ exploits this by extrapolating the predictions of the mean-field theory down from the upper critical dimension d_{u} by an asymptotic expansion in the "small" parameter $\epsilon \equiv 4 - d$, known as the $\epsilon-$expansion. This is how it is often done, technically.

The underlying intuitive picture, which is independent of the particular technical implementation, is the following. Above the upper critical dimension d_{u}, there are enough neighbors to constrain the fluctuations to the molecular scale ℓ, so that the deterministic thermodynamic theory applies at all scales much larger than ℓ. Upon crossing d_{u}, fluctuations on all scales get coupled to each other and to the microscopic scale ℓ in a subtle way. Simple dimensional analysis then fails and exponents acquire complicated non-rational values, related to each other by the scaling hypothesis and the ensuing exponent relations. The spatial structure of the order parameter fluctuations $\psi(\mathbf{r})$ corresponds to a homogeneity — in the statistical sense — of a strangely ramified self-similar or "fractal" pattern of heterogeneities, and this is the physical origin of the mathematical structure postulated in the scaling hyothesis.

The renormalization group ("RG") targets and exploits precisely this nontrivial spatial self-similarity of the order parameter fluctuations and attributes this special symmetry (a statistical dilation symmetry) to a fixed point of a semigroup transformation. A RG-transformation consists of two steps: an incremental *coarse graining* and *rescaling*. Above and below the critical point (i.e. away from the RG fixed point), the repeated application of the RG-transformation takes one to the totally disordered and ordered states corresponding to $T = 0$ and $T \to \infty$, respectively. Right at the fixed point, where the pattern of order-parameter fluctuations is fully self-similar, the pattern remains unchanged in the statistical sense. In terms of the mathematical structure of the theory, this means that, in the vicinity of the fixed point, the (generally unknown) "true" Landau–Ginzburg functional would reproduce itself, up to some sub-leading terms, with rescaled parameters under the RG-transformation. From the rescaling of the coefficients of the Hamiltonian under renormalization one could then read off the critical exponents just as one reads off the exponent γ in the ordinary Landau theory from the coefficient of the quadratic term in $\mathcal{L}(\psi)$. Above d_u, the whole procedure simply recovers the conventional homogeneity of the ordinary Landau functional, with simple rational exponents, as familiar from van-der-Waals type models. But for intermediate dimensions $d_\mathrm{l} < d < d_\mathrm{u}$, the second, more interesting type of self-similarity emerges, corresponding to a "non-trivial" new fixed point of the RG-transformation and irrational numerical values of the exponents.

Decimation transformation

To see how the RG-idea works in practice, it is useful to consider a simple example, where it is not masked by technicalities of the perturbation theory. The exactly solvable Ising chain is a suitable toy model, although it suffers from two flaws, namely its discrete state space and the associated Schottky anomaly, manifest in the strange exponential form of the critical parameter $t = e^{-\beta J}$, and the fact that $T_c = 0$, so that one cannot truly cross the critical point. The advantage is that one can straightforwardly see how renormalization works on the level of the Hamiltonian and the exact free energy, without perturbation calculations in the couplings and in the space dimension.

The starting point is a new perspective onto Eq. (124). Instead of trying to do the sum directly to obtain the free energy (which would be completely out of reach for more interesting models), one asks the following question: how does the partition sum evolve upon taking a partial trace? The partial trace is the technical realization of the above mentioned incremental coarse-graining or low-pass filtering. For the Ising chain with nearest-neighbor interactions, it can be performed exactly and is particularly simple. Due to the completeness of the basis, tracing out every second spin simply amounts to deleting all terms $\sum_{s_{2i}=\pm 1} |s_{2i}\rangle\langle s_{2i}| = \mathbf{1}$ for integer i in Eq. (124):

$$Z_N = \sum_{\{s_{2i-1}=\pm 1\}} \langle s_1|\hat{T}^2|s_3\rangle \dots \langle s_{N-1}|\hat{T}^2|s_1\rangle \ . \tag{137}$$

This defines a new Ising chain with only half as many spins and twice the lattice constant of the original chain. In other words, Z_N can be interpreted as the partition sum $Z_{\tilde{N}}$ of a chain with parameters that have been renormalized according to the rule

$$N \to \tilde{N} = N/2 \ , \qquad \ell \to \tilde{\ell} = 2\ell \,, \tag{138}$$

and with a new transfer matrix $\tilde{T} \propto T^2$ instead of T. Specializing to the field free case, $h \equiv 0$, and writing the original transfer matrix in the form

$$T = \begin{pmatrix} t^{-1} & t \\ t & t^{-1} \end{pmatrix} \quad \text{with} \quad t \equiv e^{-\beta J} \ , \tag{139}$$

one observes the following transformation rule "under renormalization":

$$T \to T^2 = \begin{pmatrix} t^{-2} + t^2 & 2 \\ 2 & t^{-2} + t^2 \end{pmatrix} = 2\tilde{t}^{-1} \begin{pmatrix} \tilde{t}^{-1} & \tilde{t} \\ \tilde{t} & \tilde{t}^{-1} \end{pmatrix} \equiv 2\tilde{t}^{-1}\tilde{T} \ . \tag{140}$$

Up to the overall factor $2\tilde{t}^{-1}$, we thus have

$$t \to \tilde{t} \quad \text{with} \quad 2\tilde{t}^{-2} \equiv t^{-2} + t^2 = 2\cosh 2\beta J \ , \tag{141}$$

for the matrix elements. This corresponds to

$$2K \to 2\tilde{K} = \ln\cosh 2K \tag{142}$$

for the dimensionless coupling strengths $K \equiv \beta J$, which shows that the renormalization steps can be thought of as causing a shift in temperature or interaction strength. In order to make the model exactly self-similar under the RG transformation, the partition sum Z_N and the corresponding free energy per spin, $f \equiv -(\ln Z_N)/(\beta N)$, have to absorb the remaining overall factors $2\tilde{t}^{-1}$ in Eq. (140) in each step. Their transformation rules thus read

$$Z_N(K) = 2^{\tilde{N}} \cosh^{\tilde{N}/2}(2K) Z_{\tilde{N}}(\tilde{K}) \, , \tag{143}$$

$$(N/\tilde{N})\beta f(K) = \beta f(\tilde{K}) - \tilde{K} - \ln 2 \ . \tag{144}$$

Using $2\tilde{K} \sim 2K - \ln 2$ for large K from Eq. (142), this can be rewritten in the form

$$(N/\tilde{N})[\beta f(K) + K] = \beta f(\tilde{K}) + \tilde{K} \tag{145}$$

which again, as in Eq. (135), identifies $f(J) + J$ as the singular part of the free energy. For a regular lattice with lattice unit ℓ in $d = 1$ space dimension, the factor $N/\tilde{N}$ can, according to Eq. (138), be written as $\tilde{\ell}^d/\ell^d$, which diverges upon repeated decimation, as does $\hat{\xi}^d$ upon approaching the critical point, in agreement with Eq. (152) and the discussion after Eq. (106).

The above recursion relations under decimation are called "renormalization group equations". They can be read as an equations of motion for the couplings K and the free energy f under coarse-graining. By solving them, one obtains the renormalization group flow. For example, for a ferromagnet with $K > 0$, any finite coupling strength is driven towards the origin ($K = 0$), corresponding to the high-temperature limit of a paramagnet, with the RG-flow given by Eq. (142) (sketch). The limits $K = \infty$ and $K = 0$ thus correspond to the unstable and stable fixed point of the renormalization group equations, describing total order and disorder of the spins, respectively. A typical chain conformation at low temperature may have large aggregates of aligned adjacent spins, but these will shrink in successive coarse-graining steps and eventually degenerate into single spins pointing in arbitrary directions. After sufficiently many coarse-graining steps, the system looks like a pure paramagnet. With a conformal mapping (in the variable t^2) to the new critical parameter

$$z \equiv \frac{1-t^2}{1+t^2} = \tanh K \leq 1 \ , \tag{146}$$

Eq. (142) takes the strikingly simple form

$$\tilde{z} = z^2 \tag{147}$$

which makes the above statements most obvious (iteration clearly takes any $z < 1$ to zero). It moreover gives direct access to the critical exponent ν if one considers $\xi(z)$ near the unstable critical point (corresponding to the phase transition). Since the decimation transformation, as a purely formal renaming exercise does of course not change the physics of the system, it leaves the physical correlation length invariant,

$$\xi \equiv \tilde{\ell}\hat{\xi}(\tilde{z}) = \ell\hat{\xi}(z) = \text{const.} \ , \tag{148}$$

but shrinks the dimensionless correlation length $\hat{\xi}$, measured in units of the lattice constant ℓ that increases by $\tilde{\ell}/\ell = 2$ during each renormalization step:

$$\hat{\xi}(\tilde{z}) = \hat{\xi}(z^2) = \hat{\xi}(z)\ell/\tilde{\ell} = \hat{\xi}(z)/2 \ . \tag{149}$$

The solution of $\hat{\xi}^{-1}(z^2) = 2\hat{\xi}^{-1}(z)$ is

$$\hat{\xi}^{-1} \propto -\ln z = -\ln \tanh K \sim 2e^{-2K} \propto t^2 \ , \tag{150}$$

where the final asymptotic result holds for large $K = \beta J$ (small t), corresponding to the low-temperature limit. This suggests the value $\nu = 2$ for the critical exponent characterizing the divergence of the correlation length ξ at low T. Via the hyperscaling relation $d\nu = 2 - \alpha$, one concludes $\alpha = 0$, consistent with a vanishing (non-diverging) specific heat, which vindicates the choice $t = e^{-\beta J}$ for the critical parameter. Together with the result $\chi = \beta e^{2\beta J}$ for the susceptibility from Eq. (128), which yields $\gamma = 2$, and the exponent relation $\gamma = \nu(2-\eta)$ deduced from Eq. (116), one finds a comparatively large anomalous dimension $\eta = 1$. And $2-\alpha = 2\beta+\gamma$ implies that the exponent β for the growth of the order parameter below T_c vanishes, $\beta = 0$, in accord with the absence of a spontaneous magnetization for any finite coupling strength.

Yang–Lee and Fisher zeros and Julia sets

The above calculations have shown explicitly how thermodynamic quantities develop non-analytic behavior as a function of the coupling strength and the particle number. The same holds for the external field. It may still appear miraculous how such singularities can come about by the apparently innocent operation of adding positive and analytical Boltzmann factors. The explanation is provided by the complex zeros of the partition sum that yield singularities if one takes the logarithm to obtain the free energy. These zeros of the partition sum (in the complex h or t plane) are known as Fisher and Yang–Lee zeros, respectively. For the former, starting from the partition sum for finite N and $\beta J \to \infty$ at finite field $h > 0$, corresponding to the free energy in Eq. (130), one has to find the complex solutions of

$$Z = 2e^{N\beta J} \cosh N\beta h = 0 \quad \Rightarrow \quad \beta h = \pm i\pi(n - 1/2)/N \quad n \in \mathbb{N}\,. \tag{151}$$

While these "Fisher zeros" lie on the imaginary βh-axis, they close up to the origin for increasing N and eventually accumulate there in the limit $N \to \infty$ — thereby destroying the analyticity at the origin by literally cutting the real axis into two halves.

Similarly, starting from the partition sum for vanishing field ($h = 0$), and using $Ne^{-2\beta J} \ll 1$ and $(1 + x/N)^N \sim e^x$ for large N, one finds

$$Z = (e^{\beta J} + e^{-\beta J})^N + (e^{\beta J} - e^{-\beta J})^N = e^{N\beta J}[(1 + e^{-2\beta J})^N + (1 - e^{-2\beta J})^N] \tag{152}$$

and thus the zeros of Z in the complex "temperature" plane from

$$Z \overset{N\to\infty}{\sim} 2e^{N\beta J} \cosh[Ne^{-2\beta J}] = 0 \quad \Rightarrow \quad t^2 = \pm i\pi(n - 1/2)/N \quad n \in \mathbb{N}\,. \tag{153}$$

As for the complex $h-$plane, there is again an accumulation of zeros that move in towards the real axis like scissors and eventually cut the complex $t-$plane into pieces at the origin for $N \to \infty$. The zeros approach on the imaginary axis of the complex t^2-plane or along the diagonals in the complex t-plane, respectively.

While the general conclusion of the accumulation of zeros at the critical point is correct, the problem with the above arguments is that they invoke the conditions $\beta h \ll 1$ and $Ne^{-\beta J} \ll 1$, which, if extended to the complex plane, are in conflict with the values of the larger zeros (which are thus not reliable). The conformal mapping Eq. (146) is helpful in this respect, and moreover convenient for establishing a connection between the zeros of the partition sum and so-called Julia sets. It maps the imaginary axis of the complex t^2-plane onto the unit circle of the complex $z-$plane. Of the four special points $t^2 = \pm i, 0, \infty$ the first two are invariant under the conformal mapping (they go to $z = \pm i$), the remaining two (corresponding to the low temperature fixed points for the ferro- and antiferromagnet, respectively) go to $z = \pm 1$. Equation (146) therefore also maps the complex zeros of the partition sum in the complex t^2-plane, which all lie on the imaginary axis, onto the unit circle in the complex $z-$plane. Starting with a ring made of two spins,

$$Z_2 = \mathrm{tr}\,\hat{T}^2 = 2(t^{-2} + t^2) = 0 \;\Rightarrow\; t^2 = \pm i \;\Rightarrow\; z = \pm i \;, \tag{154}$$

one easily finds the $z-$values corresponding to the partition-sum zeros of long Ising chains ($N = 2^m$) by applying the inverse of the decimation transformation, Eq. (147). For this procedure, it is not necessary to require t to be small. By repeatedly taking the square-root of z (i.e. simply dividing its phase angle by two), one thus finds the exact complex "Yang–Lee" zeros

$$z = (\pm i)^1, \ldots, (\pm i)^{1/2^m} = e^{\pm i\pi/2}, \ldots, e^{\pm i\pi/2^m} \;. \tag{155}$$

They accumulate densely on the real axis at the low-temperature fixed point at $z = 1$ (corresponding to $t \to 0$, $K \to \infty$, $T \to 0$, respectively), in the limit N, $m \to \infty$. And it definitely matters whether you are on the left or right of the point where the real axis is pinched, or, in other words, which limit is taken first, $N \to \infty$ or $T \to 0$.

The RG transformation maps all the zeros in Eq. (155) onto each other. They cannot leave the unit circle in the complex $z-$plane, which is invariant under the transformation. It represents the border between the points that iterate to the trivial fixed point $z = 0$ (corresponding to the paramagnet with $t^2 = 1$, $K = 0$, $T \to \infty$) under the renormalization transformation, Eq. (147), and those iterating to infinity (which have no obvious physical interpretation for the Ising chain). One can ask the general question, how such "border sets" or *Julia sets* look like for more complicated transformation rules, e.g. for $z \to z^2 + c$. This produces somewhat more impressive examples, which provide a nice and intuitively appealing hint at the (hidden) complexity of generic interacting many-body systems and their complicated but universal critical behavior. One will naturally wonder for which values of c these sets are connected and thus cut the real axis into analytically disjointed regions. This leads one to the study of the famous *Mandelbrot set* (the complex numbers c for which the iteration does not diverge when started from $z = 0$). Colorful illustrations and historical remarks can e.g. be found in the book "The Beauty of Fractals" by H.-O. Peitgen and P. H. Richter.

5 Stochastic Thermodynamics

The preceding chapter has demonstrated some procedures to get hold of the critical fluctuations that elude the usual density functional theories. Another way to approach the problem is to simply add the fluctuations by hand to the hydrodynamic theories such as Landau–Ginzburg theory or other DFTs. This yields so-called fluctuating hydrodynamic theories. For example, starting from the Euler–Lagrange equation

$$\frac{1}{n_c k_B T_c}\frac{\delta L_G}{\delta\psi} \equiv \nabla\frac{\partial\mathcal{L}_G}{\partial\nabla\psi} - \frac{\partial\mathcal{L}_G}{\partial\psi} = \ell^2\nabla^2\psi - t\psi - g\psi^3 + h = 0 \tag{156}$$

for the GL free energy, Eq. (95), one can add a random force term (in addition to or in place of h) that represents the thermal noise and drives the hydrodynamic equation. Beyond critical phenomena, which are not pursued any further here, such an approach is very useful for many things, in particular for describing the mesoscopic dynamics of many-body systems, which is neither fully deterministic, as the macroscopic behavior, nor quite as wild as the molecular chaos — and therefore often quite interesting. It is also a good starting point for dealing with fluctuations and dynamics very far from equilibrium and reveals the real power of statistical mechanics, beyond its ability to provide thermodynamic state equations. The paradigmatic case to be discussed below is a system near thermal equilibrium, with a strong scale separation between the slow modes explicitly resolved in the dynamic equation, and the fluctuating "microscopic" degrees of freedom summarily represented by the thermal noise. Close to equilibrium, there is no need to actually microscopically calculate the dynamics of the thermal forces. Their statistical properties follow from fundamental symmetries and the structure of the Gibbs ensembles. Einstein's seminal paper on Brownian motion provides a perfect illustration of this efficient and convenient approach.

5.1 Einstein 1905: Brownian motion in a nutshell

Einstein's paper[29] from 1905 nicely summarizes, in a nutshell, the various levels of a statistical mechanics descriptions of a many-body system with a scale separation between microscopic and mesoscopic degrees of freedom, namely thermostatics, thermodynamics, and stochastic thermodynamics. It exploits the "middle-world"[30] character of Brownian motion, intermediate between the molecular and the macroscopic world, to arrive, in a few lines, at the simplest versions of the corner stones of equilibrium stochastic thermodynamics.

Brownian motion, as Einstein puts it, is the thermal motion of particles suspended in a solvent, which are small enough to jiggle perceptibly, but large enough to be visible in the microscope. Einstein proposed to study this motion as a probe of the atomistic structure of the (then) invisible molecular world, i.e. to exploit that it provides a scaled-up and more tangible model of the invisible molecular world. Indeed Perrin received the Nobel prize in 1926 "for proving atoms real" on this basis, which eventually quieted protests by those calling themselves "Energetiker" (Ostwald and some other anti-atom freaks in Leipzig). Paraphrasing

---

[29] *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen*, Annalen der Physik, Leipzig 17 (1905) 549

[30] Mark Haw: Middle World, the restless heart of life and matter, MacMillan Science.

Einstein, one could also say that Brownian motion is the thermal motion of dispersed particles small enough to jiggle perceptibly, but large enough to admit some coarse-graining over the solvent degrees of freedom. In fact, Einstein actually describes the solvent in the simplest possible hydrodynamic approximation (known today as the "Markov approximation"), which greatly simplifies his task. This statement may justly prompt the question how this could ever lead to a proof of the atomic structure of matter, and a proper historian would moreover note that what is commonly called the Einstein relation, was first derived by an Australian researcher named Sutherland. But then, you should of course not mistake our common myths of scientific discovery for a proper history of science.

Step 1: thermostatic force balance

Einstein argues against the view — still common at the time — that suspended particles may be dismissed in thermodynamics. He insists that N particles, each of volume $v_\circ$, which are suspended in a solvent at small number density $n \ll v_\circ^{-1}$ represent an (almost) ideal gas and therefore give rise to an osmotic pressure[31]

$$p = nk_BT \,, \tag{157}$$

thereby exerting a finite, albeit small, thermodynamic force if spatially restricted. In presence of a constant volume force $\mathbf{K}$ (e.g. gravity, $\mathbf{K} = -mg\mathbf{e}_z$) acting on the particles one thus has the force balance

$$n(\mathbf{r})\mathbf{K} = \nabla p(\mathbf{r}) \,, \tag{158}$$

or, using Eq. (157) in Eq. (158), the barometer equation

$$p(\mathbf{r}) = n(\mathbf{r})k_BT \propto e^{\mathbf{K}\cdot\mathbf{r}/k_BT} = e^{-mgz/k_BT} \tag{159}$$

in isothermal equilibrium ($T = \text{const.}$).

Step 2: detailed balance of thermodynamic fluxes

In a second step, Einstein suggests to reconsider this stationary balance of thermodynamic forces as a *dynamic balance of the thermodynamic fluxes* excited by these forces, namely the diffusion and drift currents

$$-\mathbf{j}_{\text{diff}} = \mathbf{j}_{\text{drift}} \equiv n\mathbf{v}_{\text{drift}} \,. \tag{160}$$

He employs two (linear-response) constitutive equations that were well established by 1905 to express the currents in terms of the forces: Fick's first law of diffusion[32] (from 1855),

$$\mathbf{j}_{\text{diff}} = -D\nabla n \,, \tag{161}$$

---

[31] A truly thermostatic approach would use the gas constant and count matter in moles.

[32] Combining it with mass conservation, $\partial_t n + \nabla\cdot\mathbf{j}_{\text{diff}} = 0$, yields Fick's 2nd law, Eq. (165).

and Stokes' law for the viscous drag on a suspended particle in uniform motion (from 1851),

$$\mathbf{v}_{\text{drift}} = \mathbf{K}/\zeta \ . \tag{162}$$

The two response coefficients or susceptibilities, ζ and D, characterize the friction and diffusivity of the Brownian particles, respectively. Eliminating $\mathbf{K}$ and n between the force and flux balance equations (158) and (160), one infers the *Sutherland–Einstein relation*[33],

$$D = k_B T/\zeta \ . \tag{163}$$

This is a relation between two transport coefficients appearing in two generalized hydrodynamic equations, namely the above equations by Stokes and Fick. What made Eq. (163) historically remarkable, is the appearance of the Boltzmann constant k_B (then still called Planck's constant) originally written in the form of the gas constant divided by Avogadro's number (then called Loschmidt number). The presence of this truly microscopic quantity in Eq. (163) meant that it could be deduced from measurements of macroscopic (or at least mesoscopic) quantities, alone. Intriguingly, Eq. (163) thereby indicated a relation of the kinetic coefficients D, ζ in the hydrodynamic equations by Fick and Stokes to the microscopic, molecular world and its discrete atomistic structure (the ratio D/ζ would vanish in the continuum limit of infinite Loschmidt number). In the joints between some unsuspicious smooth continuum equations lurked, much to the dismay of Ostwald, the grotesque face of the atomic world. In this sense, Eq. (163) is reminiscent of another famous relation put forward by Einstein at about the same time, $E = \hbar\omega$, and played an equally important role for establishing the discrete atomistic structure of matter. In contrast to the latter, Eq. (163) was not postulated but derived — namely from the (at the time neither new nor generally accepted) assumption that thermodynamics applies to suspended particles. There was no reference, though, to the molecular structure of the solvent that Einstein tacitly coarse-grained away.

Step 3: stochastic thermodynamics

The explicit link to the underlying thermal fluctuations is established at the end of Einstein's paper, which undertakes a probabilistic derivation of the diffusion equation based on a particularly simple model for Brownian fluctuations, nowadays called a *random walk*. A random walk is just a caricature of the real physical Brownian motion. But Einstein appeals to the *universality* of the hydro-/thermodynamic laws, anticipating that the simplest conceivable representative of a huge universality class of microscopic systems will do as good a job as any other for his purpose.

---

[33] "Stokes–Einstein" if $\zeta = 6\pi\eta a$ for a particle of radius a in a solvent of viscosity η is used.

Einstein's starting point is the interpretation of the particle concentration $n(x,t)$ as probability density to find a single Brownian particle (for simplicity in one space dimension) at position x. Leaving its normalization for later, and using common sense, he expresses its value at time $t+\tau$ as a function of its values at time t,

$$n(x,t+\tau) = \int \mathrm{d}y \; n(x+y,t)\varphi_\tau(y) \; . \tag{164}$$

All information that is left about the solvent in this *Master–* or *Chapman–Kolmogorov*-type equation, as it would be called today, is now hidden in $\varphi_\tau(y)$. This auxiliary "jump probability" for the particle to move a distance y in time τ should be sharply peaked around $y=0$, normalized, and (for a homogeneous space) symmetric. By imposing these natural conditions on $\varphi_\tau(y)$ and Taylor expanding $n(x,t)$ with respect to τ and y, Einstein circumvents the construction of an explicit *microscopic dynamical theory* of the process. Matching the leading order terms on both sides (work it out as an easy exercise!), he recovers the diffusion equation

$$\partial_t n(x,t) = D\nabla^2 n(x,t) \tag{165}$$

along with a stochastic expression for the diffusion coefficient

$$2D = \frac{1}{\tau}\int \mathrm{d}y \; y^2\varphi_\tau(y) = \frac{\langle \Delta x^2\rangle}{\tau} = \frac{\langle \Delta \mathbf{r}^2\rangle}{\tau\, d} \; . \tag{166}$$

Einstein understands the expressions for D (with d the spatial dimension) to be independent of the time interval τ considered, as long as τ is large compared to a molecular collision time[34], despite the dubious limit "$\tau \to 0$". As suggested by Eq. (166), and as you will verify in greater detail in the exercises, a crucial point is that the mean-square displacement (MSD) per time unit τ remains a well defined quantity at least for a large range of values of τ, quite in contrast to the velocity $\langle|\Delta x|\rangle/\tau$, which some experimentalists had been trying to measure. What they naturally expected to be a "good" variable on the basis of Newton's law of motion turned out to be meaningless, which reveals the crucial importance of a good theory (or a good intuition) about what is worthwhile to be measured, in particular for phenomena governed by strong fluctuations.

Coda: Langevin's noisy hydrodynamics

Three years after Einstein's paper appeared, Langevin wrote a paper where he proposed another very elegant way of dealing with the fluctuations of Brownian particles. He simply started from the equation of motion for a particle with friction and added a random force to it, in the simplest case (if inertia is negligible):

$$\zeta\dot{x} = \xi \qquad (t>0)\,. \tag{167}$$

---

[34] Actually, τ even has to be larger than the (much longer) solvent relaxation time.

This is exactly the strategy pursued in Eq. (156), above, however not for a field equation (the GL Euler–Lagrange equation) but for Stokes' equation for a particle suspended in a solvent. Ornstein worked out the properties of the random "noise" force $\xi(t)$. Start, for example, from Einstein's final results, Eqs. (163), (166), which link the Brownian fluctuations to diffusivity, or temperature and friction:

$$2D\tau = \langle [x(\tau) - x(0)]^2 \rangle = \langle \left(\int_0^\tau \mathrm{d}t \, \dot{x}(t) \right)^2 \rangle \overset{\text{Eq. (167)}}{=} \int_0^\tau \mathrm{d}t \int_0^\tau \mathrm{d}t' \, \frac{\langle \xi(t)\xi(t') \rangle}{\zeta^2} \, . \tag{168}$$

For this to hold at all times τ, the noise needs to be $\delta-$correlated according to

$$\langle \xi(t)\xi(t') \rangle = 2k_B T \zeta \delta(t - t') \qquad \overset{\text{TTI}}{\Longrightarrow} \qquad \langle \xi(t)\xi(0) \rangle = 2k_B T \zeta \delta(t) \, . \tag{169}$$

In the second line, the time-translation invariance (TTI) or "stationarity" of equilibrium averages has been exploited to shift the origin of the time axis[35] to t'. The fact that the product of the friction coefficient and temperature quantifies the strength of the thermal noise is general and goes under the name of *fluctuation-dissipation theorem* (FDT) "of the second kind". Note that it specifies the statistical properties of the noise completely, if the latter is assumed to have no systematic contribution (i.e., vanishing mean $\langle \xi \rangle = 0$) and to be Gaussian. The Gaussian distribution is not only a technically convenient assumption. It may (loosely) be motivated by the fact that the noise represents the sum of a large number of weakly correlated molecular collision events, as familiar from the arguments put forward for justifying the canonical ensemble. Thereby, the construction of an explicit dynamical model of the solvent degrees of freedom has again (as in Einstein's discussion) been avoided.

The resulting model defined by Eqs. (167), (169), rests on two major idealizations, though, namely that memory can be neglected both in the friction and in the thermal noise. The assumption that the thermal force fluctuates wildly with no temporal correlations, and that Stokes' law, originally derived for stationary forcing, can be applied to such unruly force protocols, amounts to the so-called *Markov approximation*, which is also implicit in Einstein's derivation. It entails that the Fourier transform of Eq. (169), the so-called *"power spectrum"* (sometimes symbolically written as $\langle \xi_\omega \xi_{-\omega} \rangle$)

$$\int_{-\infty}^{\infty} \mathrm{d}t \, \langle \xi(t)\xi(0) \rangle e^{i\omega t} = 2\zeta k_B T \tag{170}$$

is flat, i.e., it contains all frequencies with equal strength, and is therefore called white noise. With this particular form of the noise correlations, the laws of Brownian motion as discussed by Einstein may completely be recovered from this intuitive model. (You may try to show this as an exercise.) In reality, the plateau in the power spectrum will extend over a finite frequency range

---

[35] First write $\langle \xi(t)\xi(t') \rangle = \langle \xi(t - t')\xi(0) \rangle$, then rename $t - t'$ as t.

only, and the $\delta-$function in Eq. (169) will have to be replaced by a symmetric function with a finite width. Its time symmetry, or time-reversal invariance (TRI), is guaranteed by the time-translation invariance of equilibrium averages, which entails time-reversal invariant auto-correlations $\langle\xi(t)\xi(0)\rangle = \langle\xi(|t|)\xi(0)\rangle$ via $\langle a(t)b(0)\rangle = \langle a(0)b(-t)\rangle = \langle b(-t)a(0)\rangle$ for $a \equiv b \equiv \xi$.

Fluctuation-Dissipation & Green-Kubo relations

Some further, equivalent formulations are easily derived from Eqs. (167), (169). Integrating Eq. (169) over the whole time axis or, equivalently, taking $\omega \to 0$ in Eq. (170) one immediately gets the following integral representation for the friction coefficient,

$$\zeta = \frac{1}{2k_BT}\int_{-\infty}^{\infty}\mathrm{d}t\,\langle\xi(t)\xi(0)\rangle \overset{\mathrm{TRI}}{=} \frac{1}{k_BT}\int_{0}^{\infty}\mathrm{d}t\,\langle\xi(t)\xi(0)\rangle = \lim_{\omega\to 0}\frac{\langle\xi_\omega\xi_{-\omega}\rangle}{2k_BT} \tag{171}$$

and, using Eqs. (163), (167), the diffusion coefficient

$$D = \int_0^\infty \mathrm{d}t\,\langle\dot{x}(t)\dot{x}(0)\rangle\,. \tag{172}$$

Here, the time-translation invariance of equilibrium averages has again been exploited to shift the origin of the time axis to t', and the time symmetry of the correlator to restrict the integration range to positive times. Such *"Green–Kubo relations"*, expressing a transport coefficient as a time integral over correlation functions of equilibrium fluctuations (or, equivalently, as the low frequency limit of their power spectral density), are generally very useful for non-invasive measurements of transport coefficients, i.e., without actually forcing a system, in numerical simulations or experiments. Also notice the formal analogy of the last expression in Eq. (171) with the so-called *sum rule* in Eq. (18), which expresses a transport coefficient (the compressibility) in terms of a spatial "low-frequency" limit of a correlation function (the structure factor).

Further, integrating Eq. (169) only from $-t$ to t, and using Eq. (167), one finds

$$2\frac{k_BT}{\zeta} = \int_{-t}^{t}\mathrm{d}t'\,\langle\dot{x}(t')\dot{x}(0)\rangle = \langle x(t)\dot{x}(0)\rangle - \langle x(-t)\dot{x}(0)\rangle\,, \tag{173}$$

and with the time asymmetry of $\langle x(t)\dot{x}(0)\rangle = -\langle x(-t)\dot{x}(0)\rangle$

$$\frac{1}{\zeta}\mathrm{sgn}(t-t') = \frac{1}{k_BT}\partial_{t'}\langle x(t)x(t')\rangle\,. \tag{174}$$

Rewriting this once more, using the causal response function or susceptibility $\langle\delta x(t)/\delta f(t')\rangle$ (that vanishes for $t < t'$) and the integrated Stokes equation

$$x(t) = \int_{-\infty}^{t}\mathrm{d}\tau\,f(\tau)/\zeta \quad\Rightarrow\quad \frac{\delta x(t)}{\delta f(t')} = \frac{1}{\zeta}\theta(t-t') \tag{175}$$

yields the *fluctuation-dissipation theorem* (FDT) "of the first kind"

$$\langle \delta x(t)/\delta f(t') \rangle = \frac{\theta(t-t')}{k_B T} \partial_{t'} \langle x(t) x(t') \rangle \, , \tag{176}$$

which is a dynamic generalization of the fluctuation response relations discussed earlier in the lecture.

There are a couple of interesting (and quite general) observations to be made, here. First, the step function $\theta(t)$ reconciles the causality of the response function with the time-reversal invariance (TRI) of the equilibrium correlation function $\langle x(t)x(t')\rangle$ and the corresponding time asymmetry of its time derivative. Secondly, the signum function in Eq. (174) suggests a sign change of the friction term in Eq. (167) upon time reversal, a property referred to as *"passivity"* (friction sucks — always, you cannot escape it, not even by running backwards in time), which is a reformulation of the second law. The discontinuous jump in the friction function is clearly a consequence of the Markov approximation of a $\delta-$correlated thermal noise. In the more general non-Markovian case, where one takes into account the persistence (or memory) of the thermal forces, the discontinuous friction function $\zeta^{-1}\text{sgn}(t-t')$ is replaced by a continuous time-asymmetric friction function $2i\chi''(t)$ defined via $2i\chi''(t-t')\theta(t-t') \equiv \langle \delta x(t)/\delta f(t')\rangle$. The FDT (in particular in the frequency domain) is then often written in the form of Eq. (174), in which case the the causal step function remains hidden in the definition of χ''.

5.2 Langevin, Ornstein, Fokker–Planck, and all that

The above crash course in stochastic (Brownian) dynamics provides a broad overview over the key results in the field, but may provoke a number of questions. The calculation of time-dependent correlation functions $\langle a(t)b(0)\rangle$ was apparently achieved without ever explicitly implementing the equilibrium averages $\langle \dots \rangle$. In fact, the symbol $\langle \dots \rangle$ acquired a new meaning (averaging over all possible realizations of the thermal noise force ξ) in the Langevin picture, and it would be desirable to understand its precise relation with the canonical definition in the Gibbs ensembles. Moreover, one would like to make the relation between the Langevin picture and the diffusion equation more explicit and to see how such equations emerge from a microscopic many-body description in terms of the Liouville equation. All this is what the present section is about.

Irreversible Thermodynamics

When setting up a Langevin equation, one practically always has some "hydrodynamic" or "thermodynamic" phenomenon in mind that is subject to some thermal noise if considered on a mesoscopic scale. The dynamics should be dissipative, since dissipation provides the link to thermal noise. To illustrate the point, one may retrace Einstein's steps, once more. Step 1 is about thermostatic force balances, a topic that was formalized in the first two chapters of this lecture. Step 2 enters dynamics in the sense of dissipative fluxes, but it simply takes over the dynamic equations from Fick and Stokes. Stokes arrived at Eq. (162) by solving the hydrodynamic equation of motion of a slowly flowing liquid (the Stokes equation), which is essentially a mechanical-engineering problem. Also Fick's equation (161) and the diffusion equation (165) emanating from it with a little help from the continuity equation,

$$\partial_t n(\mathbf{r},t) + \nabla \cdot \mathbf{j}_n(\mathbf{r},t) = 0 \ , \tag{177}$$

are such hydrodynamic equations. So how does one arrive at these hydrodynamic or thermo*dynamic* (as opposed to thermostatic) equations, in the first place? As it turns out, Fick's first law may be obtained as an immediate consequence of Stokes' mechanical law $\mathbf{v} = \mathbf{K}/\zeta$ and the static force balance[36] $n\mathbf{K} = -\nabla p$ invoked by Einstein, which allows to write the particle flux in the form

$$\mathbf{j}_n = n\mathbf{v} = -\frac{\nabla p}{\zeta} = -\frac{(\partial_n p)_T}{\zeta}\nabla n = -\frac{1}{\kappa_T n \zeta}\nabla n = -D\nabla n \ . \tag{178}$$

Fick's *gradient diffusion coefficient* can thus be expressed as

$$D = \frac{1}{\kappa_T n \zeta} = \frac{D_0}{\kappa_T n k_B T} \ . \tag{179}$$

The second equality employs the Einstein relation for the diffusion coefficient $D_0 = k_B T/\zeta$ *of a single particle.* Indeed, Fick's D is seen to reduce to D_0 (only) for a dilute suspension, for which the isothermal compressibility takes the ideal-gas form $\kappa_T = (nk_B T)^{-1}$. Generally, the gradient diffusion coefficient D in a dense suspension will thus differ from D_0 as much as the compressibility is affected by mutual particle interactions. (Actually, an improved version of this discussion would additionally account for hydrodynamic interactions between the particles, which are not contained in Stokes' law.)

Another ("more thermodynamical") way of looking at Fick's first law is obtained by the Gibbs–Duhem relation $\mathrm{d}\mu)_T = n^{-1}\mathrm{d}p$, namely

$$\mathbf{j}_n = -\frac{\nabla p}{\zeta} = -\frac{D_0 n}{k_B T}\nabla \mu \ , \tag{180}$$

which suggests gradients of the chemical potential $\mu(\mathbf{r},t)$ in place of osmotic pressure gradients as the driving force behind the phenomenon of diffusion. This alternative perspective also naturally emerges from the free energy of an isothermal isochoric system, and betrays Fick's law as a direct manifestation of the second law of thermodynamics. Namely, normalizing the free energy to a fixed volume V, and interpreting the total differential as a time derivative, one has

$$\mathrm{d}\mathfrak{f})_{T,V} = \mu\, \mathrm{d}n \qquad \Rightarrow \qquad \partial_t \mathfrak{f} = \mu\, \partial_t n \qquad \Rightarrow \qquad \partial_t \mathfrak{f} + \nabla \cdot \mu\, \mathbf{j}_n = \mathbf{j}_n \cdot \nabla \mu \ . \tag{181}$$

---

[36] Mind the minus sign: now the pressure is meant to cause (not balance) the particle flux.

The last step again made use of the continuity equation (177) for the particle density to recast the free energy balance into the form of a continuity equation with a source term. Since the divergence vanishes when the equation is integrated over the total volume of a closed system (it represents an outflux of free energy), the change of the total free energy is entirely due to this source term:

$$\dot{F}(t) = \int_V \mathrm{d}\mathbf{r}\, \partial_t \mathfrak{f}(\mathbf{r},t) = \int_V \mathrm{d}\mathbf{r}\, \mathbf{j}_n \cdot \nabla\mu \; . \tag{182}$$

Using Eq. (180), this becomes

$$\dot{F}(t) = -\int_V \mathrm{d}\mathbf{r}\, \frac{D_0 n}{k_B T} (\nabla\mu)^2 \leq 0 \; , \tag{183}$$

which says that diffusion consumes free energy and is therefore a dissipative process in the sense of the second law. Starting from a macrosopically inhomogeneous initial state, in which the system is only in a local (or partial) equilibrium, it proceeds spontaneously to bring the system towards global equilibrium. Gradients in the chemical potential or in the osmotic pressure can thus be identified as the natural thermodynamic driving forces behind the diffusion flux $\mathbf{j}_n$. Also note that the dynamic free energy $F(t)$ owes its time dependence to the dynamics of the relevant system variables, is bounded from below, and obeys $\dot{F}(t) \leq 0$. It thereby guarantees global stability and thus qualifies as a so-called Ljapunov functional.

You may still wonder how it is possible at all that equilibrium thermodynamics or equilibrium statistical mechanics admit and even describe time variations. Above, the time dependence was smuggled in via the continuity equation. Accordingly, the dynamic free energy $F(t)$ is a unique and immediate generalization of the thermostatic free energy. And the irreversible thermodynamics thus obtained is a theory for processes close to equilibrium, essentially for relaxations into equilibrium. Such a straightforward procedure is not available for systems far from equilibrium, which may never reach or not even approach equilibrium, such as systems in non-equilibrium steady states (a persistently sheared fluid) or aging systems (a fluid that vitrifies after a temperature quench).

The foregoing discussion explains the origin of the historical names “irreversible thermodynamics” or “nonequilibrium thermodynamics”. But one should actually better speak of generalized hydrodynamics or more simply thermodynamics if the latter notion had not been abused so much for something that had better been called thermostatics. In any case, the theory describes thermodynamic processes by deterministic laws, driven by (free) energy differences, as familiar from classical (continuum) mechanics. It is blind to the atomistic thermal fluctuations that are the ultimate microscopic reason for the dynamic relaxation to equilibrium and the second law. To bring them explicitly into the game, one can proceed along various ways.

Onsager regression

Within the Gibbs ensemble formalism, one may quite generally deal with time-dependent equilibrium correlation functions in the same way as with spatially varying equilibrium correlators that were introduced in the preceding chapters to generalize the (originally homogeneous) thermostatic relations. If global equilibrium is relaxed and replaced by local (in time) equilibrium, this amounts to the assumption of an adiabatic time-scale separation between the processes establishing the thermal equilibrium between the hydrodynamic variables at a certain instant in time and the slow drift of their corresponding equilibrium values that eventually establishes global equilibrium. Historically, the generalization from spatially to temporally inhomogeneous systems is known as Onsager's regression hypothesis, which is essentially equivalent to the classical FDT. In words, it states the dynamic equivalence of the relaxation of hydrodynamic variables (if not excited too far from their equilibrium values) with that of spontaneous thermal fluctuations or equilibrium correlations. The intuitive idea is that a moderate hydrodynamic excitation could have been caused by a spontaneous fluctuation, so that it should decay in the same way. Consider a variable $\hat{M}$ of vanishing mean $\langle \hat{M} \rangle = 0$ and a weak constant external field h that couples to it via a small perturbation term $-\hat{M}h$ in the Hamiltonian. The phase space density and average with and without the field are defined as ρ_h, $\langle \dots \rangle_h$ and ρ, $\langle \dots \rangle$, respectively. With this notation, the static fluctuation-response theorem, see e.g. Eq. (67), reads

$$\langle \hat{M} \rangle_h = \chi h = \langle \hat{M}\hat{M} \rangle \beta h \, . \tag{184}$$

The response is linear in the perturbing field with the linear susceptibility χ given by the correlator $\langle \hat{M}\hat{M} \rangle$ over $k_B T$. This result was originally derived by expanding the Boltzmann factor for the perturbed density to leading order in the perturbation Hamiltonian, namely

$$\rho_h \sim \rho(1 + \beta h \hat{M}) = \rho(1 + \hat{M}\langle \hat{M} \rangle_h / \langle \hat{M}\hat{M} \rangle) \, . \tag{185}$$

The funny rewriting in the last expression utilizes Eq. (184) and serves to upgrade it to dynamic processes. The formal trick goes as follows. Assume that the field h has been constant since $t = -\infty$ but is switched off at $t = 0$, i.e., $h \to h\theta(-t)$. Next, interpret all $\hat{M}$ without arguments as $\hat{M}(t = 0)$ and note that $\langle \hat{M}(t) \rangle = 0$ can be dropped, while $\langle \hat{M}(0) \rangle_h$ is a finite constant. Then, multiplication of both sides of Eq. (185) by $\hat{M}(t)$ and a phase-space integration formally yields an equality for $t \geq 0$ that can be identified with the regression theorem:

$$\frac{\langle \hat{M}(t) \rangle_{h(t)}}{\langle \hat{M} \rangle_h} = \frac{\langle \hat{M}(t)\hat{M} \rangle}{\langle \hat{M}\hat{M} \rangle} \qquad (t \geq 0) \tag{186}$$

The proper meaning of $\langle \hat{M}(t) \rangle_{h(t)}$ is

$$\langle \hat{M}(t) \rangle_{h(t)} = \langle \hat{M}(t) \rangle |_{\langle \hat{M}(0) \rangle = \langle \hat{M} \rangle_h} \, , \tag{187}$$

i.e. it denotes the free decay of $\langle \hat{M}(t) \rangle$ from its initial value $\langle \hat{M}(0) \rangle = \langle \hat{M} \rangle_h$ to zero. Hence, both sides of Eq. (186) are understood to decay from 1 to 0.

That this purely formal hocus-pocus makes sense is clarified by the explicit discussion of the special case of an Ising spin, $\hat{M} = s$. The conditional expectation value $\langle s(t) \rangle_{s(0)=\pm 1}$ of the spin at time $t \geq 0$ is easily expressed in terms of its possible values $s(0) = \pm 1$ at time $t = 0$, and so is the time-dependent two-point correlation function of the two:

$$\begin{aligned} \langle s(t) \rangle &= (1/2)\langle s(t) \rangle_{s(0)=1} + (1/2)\langle s(t) \rangle_{s(0)=-1} \\ \langle s(t)s(0) \rangle &= (1/2)\langle s(t) \rangle_{s(0)=1} - (1/2)\langle s(t) \rangle_{s(0)=-1} \; . \end{aligned} \tag{188}$$

Here, the conditional probabilities are turned into absolute probabilities through multiplication with the probabilities for the initial spin values. Without an external field, $s(0) = \pm 1$ both occur with probability 1/2. In presence of the weak field $h\theta(-t)$, these probabilities change to $e^{\pm \beta h}/(e^{\beta h} + e^{-\beta h}) \sim (1 \pm \beta h)/2$ to leading order in h, hence

$$\langle s(t) \rangle_{h(t)} = \langle s(t) \rangle_{s(0)=1}(1 + \beta h)/2 + \langle s(t) \rangle_{s(0)=-1}(1 - \beta h)/2 \; . \tag{189}$$

The interpretation of the average on the left is the same as in Eq. (187). Using $\langle s(t) \rangle = 0$ for the unconditioned spin in zero field, Eq. (189) can be written as

$$\langle s(t) \rangle_{h(t)} = \langle s(t)s(0) \rangle \beta h \; , \tag{190}$$

which, with $\langle s(0) \rangle_h = \beta h$ and $\langle s^2 \rangle = 1$, indeed corroborates Eq. (186).

Fluctuating hydrodynamics à la Ornstein–Uhlenbeck

The same conclusions may be reached in a somewhat more explicit way, along the route pursued by Langevin and Ornstein, already outlined above. It starts from the observation that hydrodynamic (thermodynamic) equations of motion describe dissipative relaxations to equilibrium. One may quite generally blame some sort of friction force for bringing all systematic motion to rest at late times, so that the equations are all of the type “initial dynamics halted by friction” and “initially provided energy dissipated to heat bath”. But things are not really totally at rest, in thermal equilibrium, they fluctuate thermally, which is the *ultima ratio* for speaking of temperature, entropy, and the second law. Friction, as a systematic manifestation of the stochastic dynamics on the molecular scale, is thus just one side of the medal. The thermal forces introduced in the Langevin equation are its other side, namely its non-systematic, random side, which was wiped under the carpet in the deterministic hydrodynamic description of the preceding paragraph. The effect of the fast molecular dynamics onto the slow (almost) deterministic thermodynamic or hydrodynamic variables can thus be decomposed into a systematic part (friction) and a random part (noise), which are firmly tied to each other by the requirement of thermal equilibrium (in the form of the FDT of the 2nd kind).

To see how these words can be turned into equations, consider a simple generic example, i.e., one that is less singular (and therefore less pathological) than Eq. (167). Let p be a hydrodynamic variable, which might in fact be the momentum of the Brownian particle that moves according to Eq. (167) at late times. The simplest possible hydrodynamic equation of motion $\dot{p} = -\gamma p$ has no other systematic force than the mandatory friction $-\gamma p$ with $\gamma = \zeta/m$. Its solution is $p(t) = p(0)e^{-\gamma t}$. As usual, the macroscopic theory is meant to predict the future from the past, corresponding to positive t. Now enter the noise ξ:

$$\dot{p}(t) = -\gamma\, p(t) + \xi(t) \qquad (t > 0)\ . \tag{191}$$

Obviously, if all systematic parts of the thermal force are contained in the friction term, the average of the remaining noise force must vanish,

$$\langle \xi \rangle = 0 \qquad (\text{“noisy noise”})\ . \tag{192}$$

Moreover, any systematic influence of the hydrodynamic past $p(t < 0)$ onto the present noise $\xi(0)$ should be prohibited if no such memory of the past is showing up in the systematic friction force in the first place,

$$\langle \xi(t)p(0) \rangle = 0 \quad (t > 0) \qquad (\text{“truly noisy noise”})\ . \tag{193}$$

This is equivalent to saying that the energy and momentum supplied to the bath are dissipated instantaneously — obviously an idealization for long time scales[37]. Together with the instantaneous (local in time) form of the friction, this statement constitutes the so-called *Markov condition.* It exploits the Brownian scale separation for degrees of freedom pertaining to objects that are large and slow compared to the atoms of the solvent or heat reservoir. At late times, the random force represents the cumulative effect of a large number of essentially stochastic (and only weakly correlated) sub-processes. One thus appeals to the central limit theorem and stipulates that $\xi(t)$ is, at any time t, a *Gaussian distributed random variable.* This assumption is very convenient, as Gaussians are specified by their first two moments and can always be integrated. It has proven to work very well for a large number of applications close to equilibrium but failed spectacularly for financial markets and other phenomena very far from equilibrium.

An important side effect of the introduction of the noise is the possibility to reinterpret the averages $\langle \dots \rangle$ in terms of the ensemble of realizations of the stochastic process $\xi(t)$. While this allows for generalizations of the averaging procedure far beyond the one of standard equilibrium statistical mechanics (one could for example include external driving forces or even some internal “activity”

---

[37] In reality, momentum is locally conserved in the solvent, so that it has to be dispersed by (fast) propagating sound modes and (slow) diffusive shear modes, which takes some time.

of the particles themselves, which would prevent the system from ever relaxing to equilibrium), it is of course most interesting to find the conditions under which the two are consistent. This is the next task.

The first requirement is that the known deterministic hydrodynamic equations should emerge on average. Note that this is an immediate consequence of Eqs. (191), (192). The second requirement is more subtle. Equilibrium averages must be time-translation invariant or *stationary* in the same way as spatial averages are translation invariant in spatially homogeneous systems. This implies, for example, $\langle p(t)p(t')\rangle = \langle p(t-t')p(0)\rangle$ and

$$\langle p(t)p(0)\rangle = \langle p(0)p(-t)\rangle = \langle p(-t)p(0)\rangle = \langle p(|t|)p(0)\rangle \qquad \text{(TRI)} \;, \tag{194}$$

with the last expression clearly exposing time-reversal invariance or "time symmetry" of auto-correlation functions as an immediate consequence of stationarity. The time symmetry of the auto-correlation $\langle p(t)p(0)\rangle$, in turn, implies the time-asymmetry of its time derivative,

$$\langle \dot p(t)p(0)\rangle = -\langle \dot p(-t)p(0)\rangle = -\langle p(t)\dot p(0)\rangle \;. \tag{195}$$

as promptly checked by renaming $t \to -t$ (hence $\partial_t \to -\partial_t$). Now, multiplying the Langevin equation (191) by $p(0)$ and averaging, one finds with Eq. (193)

$$\langle \dot p(t)p(0)\rangle = -\gamma\langle p(t)p(0)\rangle \qquad (t>0) \;. \tag{196}$$

The correlator obeys the deterministic hydrodynamic equation. Since the left side is symmetric and the right side is asymmetric under time reversal, an extension to negative times is only possible if one side is multiplied by the sign-function,

$$\langle \dot p(t)p(0)\rangle = -\gamma \operatorname{sgn} t\,\langle p(t)p(0)\rangle \qquad (\forall t) \;. \tag{197}$$

The obvious solution is $\langle \dot p(t)p(0)\rangle = \langle p^2\rangle e^{-\gamma|t|}$, which is now perfectly acceptable for an equilibrium correlation function and establishes the link between TTI and the time asymmetry of the friction function. Take another time-derivative to find

$$\langle \dot p(t)\dot p(0)\rangle = 2\gamma\,\delta(t)\langle p^2(0)\rangle + \gamma\operatorname{sgn}(t)\langle \dot p(t)p(0)\rangle \;. \tag{198}$$

Hence, altogether, the proper extension of Eq. (191) to negative times must be[38]

$$\begin{gathered} \dot p(t) = -\gamma\operatorname{sgn}(t)p(t) + \xi(t) \quad \Leftrightarrow \quad \partial_{|t|}p = -\gamma p + \xi \\ \langle \xi(t)\rangle = 0 \;, \qquad \langle \xi(t)\xi(0)\rangle = 2\gamma\,\langle p^2\rangle\delta(t) = 2\zeta k_B T\delta(t) \;. \end{gathered} \tag{199}$$

Assuming that p denotes the momentum of a Brownian particle, the equipartition theorem and the definition $\gamma = \zeta/m$ were used, in the last step. Note

---

[38] To evaluate the noise correlation, use the equation of motion with Eq. (198) and TTI.

that the friction function is now indeed found to be time asymmetric, as already anticipated above. For Gaussian noise, which is completely specified by its first two moments (average and variance), the model definition is complete and bears the name Ornstein–Uhlenbeck process. In the exercises you check that the model reproduces the static averages of the canonical ensemble and moreover allows for a quick and easy calculation of equilibrium time-dependent correlation functions, and that even non-equilibrium averages and correlation functions can be calculated by imposing non-equilibrium initial conditions. You discover that the solutions of Eq. (199) are stationary for very long times (or for canonically averaged initial data). In fact, up to trivial variations, the Ornstein–Uhlenbeck process is the only stationary Gaussian Markov process, which makes it quite a celebrity.

The last Eq. (199) is again a formulation of the FDT of the 2$^{\text{nd}}$ kind. Onsager regression, which is a formulation of the FDT of the 1$^{\text{st}}$ kind, follows from it in the same way as already illustrated for the overdamped limit (which amounts to dropping the inertial contribution $\dot{p}$), in the preceding section. The comparison of the preceding two paragraphs underscores again that, on the Markov level, fluctuations may be described either explicitly by the Langevin formalism for some stochastic variables, or implicitly by hydrodynamic equations for their correlation functions. This link is considered in more detail in the next two paragraphs.

The Fokker–Planck equation

The Gaussian property of the noise distribution is not destroyed by a linear coordinate transformation. It is therefore transmitted from the noise correlations that drive the dynamics to the solutions of the linear Langevin equation (199). Therefore, the conditional distribution $\rho(p,t|p_0)$ to find values of the variable p at time t, given an initial condition $p = p_0$ at time $t = 0$, must be Gaussian at all times. One can thus immediately write down the general form of the solution:

$$\rho(p,t|p_0) = \frac{1}{[2\pi\Delta(t)]^{1/2}} e^{-[p-\bar{p}(t)]^2/2\Delta(t)} \; . \tag{200}$$

The explicit expressions for the dynamic first and second moments,

$$\bar{p}(t) \equiv \langle p(t)\rangle_{p_0} = p_0 e^{-\gamma|t|} \tag{201}$$

$$\Delta(t) \equiv \langle [p(t)-\bar{p}(t)]^2\rangle_{p_0} = \langle p^2(t)\rangle_{p_0} - \bar{p}^2(t) = mk_BT(1-e^{-2\gamma|t|}) \; , \tag{202}$$

follow from Eq. (199). The deterministic initial condition is recovered at short times but forgotten at late times, where the conventional canonical equilibrium distribution is approached:

$$\begin{aligned} \lim_{t\to 0} \rho(p,t|p_0) &= \delta(p-p_0) \; , \\ \lim_{t\to\infty} \rho(p,t|p_0) &= (2\pi m k_B T)^{-1/2} e^{-p^2/2mk_BT} \; . \end{aligned} \tag{203}$$

An average over the initial condition p_0 would turn the conditional distribution into the momentum distribution $\rho(p,t)$. With a Maxwellian initial momentum distribution, the latter is thus also Gaussian, otherwise it still becomes Gaussian at late times.

It is an interesting question how to get the distribution function in cases when it is not Gaussian and possibly a much richer object than suggested by its first two moments, e.g., for a nonlinear Markov processes with a more general drift term $v(p)$ in place of $-\gamma p$. The answer is provided by an equation of motion for the distribution function that can be derived for any Markovian Langenvin equation, the Fokker–Planck equation. The task is essentially to go from the "Heisenberg" picture adopted in the Langevin equation to the corresponding "Liouville" or "von Neumann" picture. For the very simple Eq. (167), Einstein's derivation of the diffusion equation (165) already achieves this task. If a (possibly nonlinear) drift term $v(p)$ is included in the Langevin equation, also the diffusion equation acquires an extra drift term. The general derivation can be accomplished with the help of the functional calculus, introduced earlier in the lecture. Here we take a shortcut, instead, sticking to the above simple example of the linear Ornstein–Uhlenbeck process, for which the *Fokker–Planck equation* for $\rho(p,t|p_0)$ is easily reverse-engineered from its known solution, Eq. (200), by differentiation:

$$\partial_t\rho = \gamma\partial_p(p\rho) + \zeta k_B T\partial_p^2\rho \qquad \text{or} \qquad \partial_{\gamma t}\rho = \partial_v(v\rho) + \frac{k_B T}{m}\partial_v^2\rho \, . \tag{204}$$

This structure survives in the case of vector variables $\mathbf{p}$ and nonlinear Markovian Langevin equations,

$$\partial_t\rho = -\partial_{\mathbf{p}}\mathbf{v}(\mathbf{p})\rho + \zeta k_B T\partial_{\mathbf{p}}^2\rho \, . \tag{205}$$

As a concrete example, consider a Brownian particle in an external force field $f(x) = -\partial_x\mathcal{U}(x)$, with the vector (x,p) generalizing the momentum coordinate p and $\mathbf{v}(p,x) = (p/m, f(x) - \gamma p)$ generalizing the drift term $-\gamma p$. For the harmonic force $f(x) = -kx$, this is a Brownian harmonic oscillator. Its Fokker–Planck equation for the density $\rho(x,p,t)$ reads

$$\partial_t\rho + \partial_x(p\rho/m) + \partial_p[f(x) - \gamma p]\rho = \zeta k_B T\partial_p^2\rho \, . \tag{206}$$

One recognizes the Liouville structure of a continuity equation $\partial_t\rho + \mathrm{div}(\mathbf{v}\rho) = 0$ for the trajectories in phase space, streaming with velocity $\mathbf{v} = (p/m, f)$ albeit with some of the momentum-space "velocity" f destroyed by friction $-\gamma p$, and the non-streaming but mass-conserving diffusion term representing the Langevin noise, on the right. Its convex and concave regions act as sinks and sources for the probabilty density, respectively, acting to erase any spatial modulations in the distribution.

In summary, the dynamic evolution of the probability distribution of fluctuating hydrodynamics variables (i.e., variables described by Markovian stochastic equations) is itself governed by a deterministic hydrodynamic equation.

The Smoluchowski equation

The Fokker–Planck operator is generally not self-adjoined, so that it is difficult to say much about its solutions, in general. The situation is much better for its overdamped version, which is obtained from the corresponding Langevin equation after dropping the inertial term $\dot{p}$ against the friction, i.e.,

$$\cancel{\dot{p}} + \zeta \dot{x} = f(x) + \xi(t) \; . \tag{207}$$

This matters so much that the corresponding Fokker–Planck equation for $\rho(x,t)$,

$$\partial_t \rho = \partial_x D(\partial_x - \beta f)\rho \quad \text{or} \quad \zeta \partial_t \rho = \partial_x (k_B T \partial_x - f)\rho \; , \tag{208}$$

deserves its own name and is called the Smoluchowski equation. For vanishing force, $f(x) \equiv 0$, it reduces to the simple diffusion equation of a single particle. Its straightforward extension to many (10^{14}, say) mutually interacting colloidal particles — with inter-particle forces $\mathbf{f}_i(\mathbf{x}) = -\partial_i \mathcal{V}(\{\mathbf{x}_j\})$ — is still a very challenging equation, though. The variable x becomes a high-dimensional vector $\{\mathbf{x}_j\})$ containing all particle coordinates and the diffusion coefficient turns into a matrix $\mathbf{D}_{ij}(\{\mathbf{x}_i\})$ that still depends on the mutual hydrodynamic interactions[39] between the particles and hence all their positions. The full equation for the $N-$particle density $\rho(\{\mathbf{x}_i\}, t)$ then reads

$$\partial_t \rho = \sum_{ij} \partial_i \cdot \mathbf{D}_{ij} \cdot [\partial_j - \beta \mathbf{f}_j]\rho \; . \tag{209}$$

It is the standard microscopic description for a wealth of problems in colloid and polymer physics.

Going back to the single-particle case with an external potential $\mathcal{U}(x)$, note that the equation can be cast into the form of a continuity equation

$$\partial_t \rho + \partial_x j = 0 \tag{210}$$

for $\rho(x,t)$, which is in then nothing but the ordinary 1-point particle density $n(x,t)$. The current

$$j(x,t) \equiv \rho v = -\zeta^{-1}(k_B T \partial_x - f)\rho \tag{211}$$

of the conserved density flows with the velocity

$$v(x,t) \equiv -\zeta^{-1}\partial_x (k_B T \ln \rho + \mathcal{U}) = -\zeta^{-1}\partial_x \mu(x,t) \; , \tag{212}$$

and is seen to be driven by gradients of the generalized chemical potential

$$\mu(x,t) = k_B T \ln \rho + \mathcal{U} = \mu_{\text{ig}}(x,t) + \mathcal{U}(x) \; . \tag{213}$$

---

[39] In the Markovian approximation, these are idealized as an instantaneous action at a distance with a complicated spatial structure that follows from the stationary Stokes equation.

The generalized chemical potential is thus (again) recognized as a neat way of defining a thermodynamic force $-\partial_x \mu$ that can be treated on par with real (mechanical, electrical, …) forces. Thanks to the mass (or probability) conserving property of the diffusive source term the thermal fluctuations can thereby smoothly be integrated into a (continuum) mechanical formalism in the form of the entropy $-k_B \ln \rho$, so that the Liouville structure of a streaming conserved phase space density is restored.

Note that the generalized chemical potential has the form of a dynamic free energy $U - TS$. More precisely, the chemical potential density $\mu\rho$ integrated over the configuration space is the excess free energy over the equilibrium free energy F_0, which governs the relaxation from a non-equilibrium initial condition to the global equilibrium state $\rho_0(x) = e^{-\beta\mathcal{U}(x)+\beta F_0}$. Exactly as in the above discussion of irreversible thermodynamics, it is the Ljapunov functional

$$\Delta F(t) = k_B T \int \mathrm{d}x\, \rho \ln(\rho/\rho_0) = \int \mathrm{d}x\, \mu\rho - F_0 = \int \mathrm{d}x\, (k_B T \ln \rho + \mathcal{U})\rho - F_0 \quad (214)$$

generalizing the equilibrium free energy that guarantees the global stability of the equilibrium state. Because of the minimum condition for the equilibrium free energy, $\Delta F(t)$ is non-negative, whereas its time derivative (at constant T and V) cannot be positive,

$$\Delta \dot{F}(t) = \int \mathrm{d}x\, \mu \partial_t \rho = -\int \mathrm{d}x\, \mu \partial_x j = \int \mathrm{d}x\, j \partial_x \mu = -\zeta^{-1} \int \mathrm{d}x\, (\partial_x \mu)^2 \rho \leq 0 \ . \quad (215)$$

For the first equality, one makes use of the normalization of the density and of the fact that the potential $\mathcal{U}(x)$ and the equilibrium free energy F_0 are conservative potentials, for the second and fourth of Eqs. (210), (212), respectively, and for the third one of the fact that nothing flows across the system boundaries.

The Smoluchowski equation thus describes fully irreversible dynamics and inherits its stability from Eq. (183). This special feature is not shared by a general Fokker-Planck or Langevin equation, nor by a Smoluchowski equation driven by an external (or internal) non-conservative dynamic field. Although they can also still be recast in the form of a Liouville continuity equation for the phase space density, they admit much richer dynamics than monotonic relaxations into equilibrium, dynamics violating Eqs. (215). A particularly important example is provided by so-called non-equilibrium stationary states (NESS), defined by the vanishing of the time derivatives in the theory (i.e. $\partial_t \rho \equiv 0$), which implies that the divergence of the current vanishes but not (necessarily) the current itself. A colloidal suspension constantly sheared at fixed shear rate is an example.

Without non-conservative driving forces, Eqs. (214), (215) assert that the long-time limit of solutions of the Smoluchowski equation is always the equilibrium state, where the current itself vanishes. In view of such perfect conduct it should not come as a surprise that the Smoluchowski–operator $\mathcal{L}_S$ with only conservative potential forces can be made self-adjoined. This is achieved by a simple normalization of the distribution ρ with the square root of its equilibrium form $\rho_0(x) \propto e^{-\beta\mathcal{U}(x)}$ (the proportionality factor is irrelevant), namely

$$\psi(x,t) \equiv \rho e^{\beta\mathcal{U}/2} , \tag{216}$$

as one might have guessed from the following rewriting of the Smoluchowski equation

$$\partial_t\rho \equiv \mathcal{L}_S\rho = D\partial_x(\partial_x + \beta\mathcal{U}')\rho = D\partial_x e^{-\beta\mathcal{U}}\partial_x e^{\beta\mathcal{U}}\rho , \tag{217}$$

which suggests to define the tilted derivative $e^{-\beta\mathcal{U}/2}\partial_x e^{\beta\mathcal{U}/2}$. Indeed, the dynamics for the normalized density ψ is straightforwardly shown to be given by

$$\partial_t\psi = De^{\beta\mathcal{U}/2}\partial_x e^{-\beta\mathcal{U}/2}e^{-\beta\mathcal{U}/2}\partial_x e^{\beta\mathcal{U}/2}\psi , \tag{218}$$

with a manifestly self-adjoined dynamic operator. Moreover, with yet some more rewriting, this equation can be recast into the form of a supersymmetric Schrödinger equation in the imaginary time variable $\tau = -i\hbar Dt$,

$$i\hbar\partial_\tau\psi = (-\partial_x^2 + \mathcal{U}_{\text{susy}})\psi , \qquad \mathcal{U}_{\text{susy}} \equiv (\beta\mathcal{U}'/2)^2 - \beta\mathcal{U}''/2 . \tag{219}$$

This establishes a 1:1 relation between solutions of the 1-particle Smoluchowski equation and the Schrödinger equation for supersymmetric potentials $\mathcal{U}_{\text{susy}}$. Since the Schrödinger equation is a subject for another lecture, and the supersymmetric form of the potential implies that every solution will find a supersymmetric partner, this seems to be a suitable point for a happy end.